31938

National Library of Canada
Bibliothèque nationale du Canada

CANADIAN THESES ON MICROFICHE
THÈSES CANADIENNES SUR MICROFICHE

NAME OF AUTHOR/*NOM DE L'AUTEUR* John Beamish

TITLE OF THESIS/*TITRE DE LA THÈSE* Topics in Non-equilibrium Physics

UNIVERSITY/*UNIVERSITÉ* University of Alberta

DEGREE FOR WHICH THESIS WAS PRESENTED/*GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* M. Sc. in Theoretical Physics

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE GRADE* 1977

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* Dr. H. J. Kreuzer

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/*DATÉ* Mar. 25 / 77 SIGNED/*SIGNÉ* John Beamish

PERMANENT ADDRESS/*RÉSIDENCE FIXE*
#105 10720 83 Ave.
EDMONTON, Alberta

National Library of Canada

Cataloguing Branch
Canadian Theses Division

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada

Direction du catalogage
Division des thèses canadiennes

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (3/77)

THE UNIVERSITY OF ALBERTA

TOPICS IN NON-EQUILIBRIUM PHYSICS

by

© JOHN BEAMISH

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE

IN

THEORETICAL PHYSICS

DEPARTMENT OF PHYSICS

EDMONTON, ALBERTA
SPRING, 1977

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled TOPICS IN NON-EQUILIBRIUM PHYSICS submitted by JOHN BEAMISH in partial fulfilment of the requirements for the degree of Master of Science in Theoretical Physics.

Supervisor

M. Razavy

Date Feb. 25, 1977

ABSTRACT

This thesis is based upon lecture notes from a course non-equilibrium physics taught by Dr. H. J. Kreuzer from 1974-1976 at the University of Alberta. My choice of topics reflects the emphasis which he put on various aspects of the theory in his course.

Part One is concerned with the thermodynamics of irreversible processes. The well-established linear theory is studied in some detail in Chapters Two and Three in which balance equations for mechanical and thermodynamic quantities are determined and the linear phenomenological laws relating "forces" to "fluxes" are studied. A fairly complete discussion of the Onsager reciprocity relations is given here. Chapter Four discusses the more recent stability theory of thermodynamic systems and its relationship to fluctuations, following the methods of Glansdorff and Prigogine. The next two chapters provide explicit examples of systems which can exhibit instabilities. Chapter Five deals with chemically reacting systems and the stability of steady states is introduced as is the concept of a limit cycle. Chapter Six discusses the Bénard problem in which instability results from the heating of a layer of liquid. The results of applying the stability criteria introduced in Chapter Four are compared to the exact result. Finally,

in Chapter Seven, the problem of a turbulent fluid is discussed as an example of a system far from equilibrium.

Part Two of the thesis discusses the microscopic foundations of macroscopic theories of irreversible processes. After some historical and introductory remarks in Chapter Eight, Chapter Nine presents some of the results and methods of classical non-equilibrium statistical mechanics, including a derivation of Boltzmann's equation and a discussion of more general kinetic equations. Chapter Ten then goes on to study the fundamental problem of irreversibility from both the classical and the quantum mechanical points of view. A discussion of the various H-theorems is given in this chapter. The quantum mechanical linear response theory is presented and analysed in Chapter Eleven, including derivations of Onsager relations and a fluctuation dissipation theorem. Finally, the last chapter deals briefly with the general theory of irreversible processes, as expounded by Prigogine and his co-workers in the years since 1957.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Dr. H. J. Kreuzer, for his excellent guidance, for his helpfulness, and for keeping me working at a more or less steady pace. This thesis is based on a set of lecture notes from a course which he taught and it strongly reflects his influence.

I would also like to extend my sincere appreciation to Mrs. Lee Cech for typing a rather long and sometimes difficult manuscript.

TABLE OF CONTENTS

PART I

Chapter		Page
I	Introduction to Part I	1
II	Balance Equations of Thermodynamics of Processes	8
	§2.1 Balance Equations for Mechanical Quantities	8
	§2.2 Models of Real Materials	22
	§2.3 Entropy Balance and the Second Law of Thermodynamics	27
III	Linear Phenomenological Laws	33
	§3.1 Choosing "Forces" and "Fluxes"	33
	§3.2 The Onsager Reciprocal Relations	39
	§3.3 Minimum Entropy Production	65
IV	Stability and Fluctuations	71
	§4.1 Liapounoff Functions and Stability	71
	§4.2 Stability of Equilibrium States	73
	§4.3 Stability of Non-equilibrium States	80
	§4.4 Stability and Fluctuations	89
V	Chemical Reactions	96
	§5.1 Equations of Motion and Stability	96
	§5.2 Examples	98
VI	Stability in Fluids at Rest (Bénard Problem)	111
	§6.1 Statement of Problem	111
	§6.2 Stability Criterion of Glansdorff and Prigogine	118
	§6.3 Nature of the Unstable Motion	124

Chapter Page

VII Stability of Fluids in Motion (Turbulence) 126

§7.1 Nature of Turbulent Flows 126

§7.2 Analytic Description of Turbulence 131

PART II

VIII Introduction and Foundations 149

§8.1 Historical Introduction 149

§8.2 Classical Foundations 152

§8.3 Quantum Mechanical Formulation 159

IX Classical Statistical Mechanics and Kinetic Theory 167

§9.1 Dynamical Equations 167

§9.2 Vlasov and Boltzmann Equations 173

§9.3 Formal Methods and Expansions 182

§9.4 Microscopic Derivation of Balance Equations 194

X Irreversibility and H-theorems 202

§10.1 Boltzmann's Equation and Irreversibility 202

§10.2 A Simple Model and General H-theorems 208

§10.3 Quantum Systems and Master Equations 227

Chapter | Page

XI Response of Quantum Systems to Mechanical Perturbations 238

§11.1 Response Theory Formalism 238

§11.2 Validity of Response Theory 243

§11.3 Applications 252

XII Prigogine Theory of Irreversible Processes 262

§12.1 Introduction and Potential Scattering 262

§12.2 General Theory and Resolvent Formalism 272

Bibliography 294

LIST OF FIGURES

Figure		Page
2.1	One-dimensional Maxwell model for viscoelasticity	26
2.2	One-dimensional Kelvin model for viscoelasticity	26
4.1	Temperature fluctuations about equilibrium	90
4.2	Temperature fluctuations about steady state	90
5.1	γ versus n_x in a reacting system	103
5.2	Solutions of eqn. (5.39)	107
5.3	A system above critical affinity	109
6.1	A possible evolution and its stability	121
7.1	Typical eddy spectrum	134
7.2	Turbulent spectrum and energy transfer	144
12.1a	A typical correlation diagram	275
12.1b	Above Fig. 12.1a in standard scattering notation	275
12.2	Six "elementary" diagrams	276
12.3	Contour for integral (12.46)	278
12.4	Lowest order vacuum-vacuum diagram	281
12.5	Multiple collision diagram	281

PART I

CHAPTER I

INTRODUCTION TO PART I

According to Planck, physical processes may be divided into three groups. First, there are the reversible processes. These are the domain of equilibrium thermodynamics and, although they form a simple and valuable class of processes, few real processes are truly reversible.

The second group he referred to as natural processes. These are the processes which tend toward equilibrium monotonically. Such processes are described by a *H*-theorem and include many non-equilibrium situations.

The final group Planck called unnatural and it included all those non-equilibrium processes which evolve away from equilibrium. In Planck's time, this last group was not considered to be very important, since it was believed that a *H*-theorem of some sort governed most physical systems. However, it is true that some systems never reach equilibrium due to the presence of some external constraint such as a temperature gradient. Also, it has since been realized that there are physical systems which do not return to equilibrium, or do so in an oscillatory manner. Examples of oscillatory systems include spin echo experiments and some types of chemical reactions. Other examples of unnatural processes are non-equilibrium phase transitions such as the transition from laminar to turbulent fluid flow, shock waves in fluids, and plastic flow. These unnatural processes usually involve non-linear effects.

A more modern division of physical processes can be made using the knowledge gained from statistical physics. First, there are the quasi-equilibrium phenomena which include the above reversible processes, as well as small fluctuations.

Secondly, there are the near equilibrium phenomena which are described by linear laws. In the macroscopic domain, this includes all of linear thermodynamics of irreversible processes and in the microscopic domain, linear response theory.

Thirdly, there are the complex and less well understood far from equilibrium processes. These include non-linear effects, instabilities, and the unnatural processes above. An important class of processes which may be in any of the above categories is the class of time independent, steady state situations, including many transport processes.

In the first portion of this thesis, the generalization of equilibrium thermodynamics to the second group of phenomena, the near equilibrium processes, will be considered. In order to get from equilibrium to non-equilibrium thermodynamics, first assume that the macroscopic system can be divided into many small cells. These cells are to be macroscopically small compared to the range over which the system's properties vary, but microscopically large so that enough particles are in each cell for the statistical method to be applicable.

Secondly, assume that the system is close enough to equilibrium that the above cells are in "local equilibrium". That is, assume that it is possible to find such cells, large enough to be considered as thermodynamic systems, which are close enough to equilibrium that thermodynamic variables such as pressure, density, temperature, and entropy can be defined as constants in each cell, varying from cell to cell.

Thirdly, treat these local thermodynamic functions as field variables and make them continuous. That is, define $T(\vec{r},t)$, $S(\vec{r},t)$, $\rho(\vec{r},t)$, etc.

Finally, write balance equations for the thermodynamic fields, giving a local description of the system.

The assumption of local equilibrium essentially is that it is possible to define a local entropy density S which is the same function of the local thermodynamic densities as the equilibrium entropy density is of the equilibrium thermodynamic densities. This implies that the Gibbs-Duhem relation is valid:

$$TdS = dU + pdV - \sum_{\gamma=1}^{N} \mu_\gamma dN_\gamma \quad . \qquad (1.1)$$

This condition, including the implicit assumption that the thermodynamic functions can be defined locally, is a macroscopic validity criterion for the extension from equilibrium to non-equilibrium thermodynamics. To explicitly calculate the conditions under which the assumption of local equilibrium is valid, the methods of non-equilibrium statistical mechanics are required, as well as a microscopic model. The assumption of local equilibrium also clearly limits one to consideration of systems which can be described by an equation of state which is independent of the gradients of the local fields. An analogy can be drawn to a dilute gas in an expanding container (cf. Glansdorff and Prigogine [1] p. 14). This is a non-homogeneous and non-equilibrium situation. However, if the expansion is not too rapid and if the system is not too dilute, the equation of state pV = RT is valid locally. Similarly, for local equilibrium, the Gibbs-Duhem relation holds.

Microscopically, local equilibrium can be understood through the Boltzmann approach, in which an arbitrary thermodynamic function $\phi(t)$ is considered to be an average of a dynamic quantity over a long time. In equilibrium,

$$\phi^{eq}(t) \equiv \lim_{\tau\to\infty} \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \phi(t')dt' \quad . \qquad (1.2)$$

The usual thermodynamic functions are independent of t in equilibrium. For non-equilibrium processes, in order for $\phi^{ne}(t)$ to be meaningful, $\phi^{ne}(t)$ can be defined as a finite time average,

$$\phi^{ne}(t) = \frac{1}{2\tau} \int_{t-\tau}^{t+\tau} \phi(t')dt' \qquad (1.3)$$

Thus, in order for $\phi^{ne}(t)$ to have the desired interpretation as a non-equilibrium thermodynamic function, instead of taking $\tau\to\infty$, as in equilibrium, it is necessary that $\tau_{rel} << \tau << \tau_{proc}$. That is, τ must be effectively infinite compared to the microscopic relaxation time τ_{rel}, but infinitesimal compared to the time scale for macroscopic thermodynamic processes, τ_{proc}. Thus it is required that $\tau_{rel} << \tau_{proc}$, which implies that local equilibrium holds only for systems where strong dissipative forces are acting to damp any fluctuations. For example, this might be the case in liquids or in moderately dense gases, but would not be true in very dilute gases where there are few collisions or in fermion systems at low temperature, where dissipative processes are ineffective.

From a practical point of view, such an approach is not useful since, in order to apply it, the exact dynamical behavior of $\phi(t)$ must be known. Therefore, as in equilibrium statistical mechanics, it is necessary to proceed through an ergodic hypothesis to ensembles and partition functions. The assumption of local equilibrium then requires that the non-equilibrium situation must be describable by a partition function close to an equilibrium partition function.

The question of the existence of some of the equilibrium functions such as entropy or temperature in systems far from equilibrium has been raised. Hofelich [2] has shown that it is possible to find a Liapounoff function which has the usual properties of entropy and which has these properties throughout the entire range where the kinetic equations of the system are valid. However, the usual properties of entropy are not enough to define such an entropy function uniquely and the usefulness of such a function is questionable.

Alternatively, Meixner [3] has discussed the possibility of a non-equilibrium thermodynamics without a local entropy. His theory involves only the exact equilibrium thermodynamic functions and the boundary conditions on the surface of the system. His approach also assumes the existence of a non-equilibrium local temperature, but does not require an operational definition of such a temperature. Meixner finds a new, integral formulation of the usual Clausius inequality $\rho \frac{ds}{dt} + \vec{\nabla} \cdot s \geqslant 0$, which does not require the existence of a local entropy density s. He then finds that the usual entropy balance equation is an equation for the entropy density exactly in equilibrium, where the entropy production depends upon the equilibrium properties, on the non-equilibrium temperature, and on the dynamic properties of the system.

Meixner's theory and the usual local equilibrium theory have the common feature that, in order for these non-equilibrium thermodynamic theories to be logically consistent, they must invoke equilibrium thermodynamics. This feature is due to the fact that such a non-equilibrium thermodynamics is an attempt to generalize an equilibrium theory to non-equilibrium situations without invoking dynamics! In other macroscopic theories, first a dynamic theory is set up and then the equilibrium theory is derived as a limiting case. It is obvious that a general non-equilibrium description must involve the dynamics of the system and thus cannot be reached via equilibrium thermodynamics.

However, there exists an essentially complete classical theory, based upon the local equilibrium assumption which will be presented in the next chapter.

CHAPTER II

BALANCE EQUATIONS OF IRREVERSIBLE THERMODYNAMICS

§2.1 Balance Equations for Mechanical Quantities

Consider a macroscopic system in a volume V, bounded by a closed surface Σ (cf. de Groot and Mazur[4]). In order to derive equations describing the time evolution of various extensive thermodynamic functions in V and fluxes through the surface Σ , it is useful to write any extensive quantity F(t) as

$$F(t) = \int_V \rho(\vec{r},t)\, f(\vec{r},t)\,dV \qquad (2.1)$$

where $f(\vec{r},t)$ is an intensive variable representing the density of F(t) per unit mass (specific variable) and $\rho(\vec{r},t)$ is the mass density as a function of position and time. Since the quantity F(t) is not necessarily conserved, the most general equation for F(t) is of the form

$$\frac{dF(t)}{dt} = P[F] + \phi[F] \qquad (2.2)$$

where P[F] is a source term describing production of F in the volume V and $\phi[F]$ is a flow term describing the exchange of F between V and its surroundings through the surface Σ . These source and flow terms may be written as

$$P[F] = \int_V \sigma_F dV$$

$$\phi[F] = -\int_\Sigma \vec{J}_F \cdot d\vec{\Sigma} \qquad (2.3)$$

where σ_F is the source or sink density in V and $\vec{J}_F$ is the current flow through Σ .

To cast this as a local equation, the Euler coordinate system is used. The Euler coordinates are fixed in space as opposed to the Lagrange system which moves with the material. Then, $\frac{dg(\vec{r},t)}{dt} = \frac{\partial g}{\partial t}$, so

$$\frac{dF(t)}{dt} = \frac{d}{dt}\int_V \rho(\vec{r},t)\ f(\vec{r},t)dV = \int_V \frac{\partial}{\partial t}(\rho f)dV$$

$$= \int_V \sigma_F dV - \int_\Sigma \vec{J}_F \cdot d\vec{\Sigma} \quad . \qquad (2.4)$$

Applying the divergence theorem to the continuous function $\vec{J}_F$ gives

$$\frac{\partial(\rho f)}{\partial t} + \vec{\nabla} \cdot \vec{J}_F = \sigma_F \quad . \qquad (2.5)$$

For mass transport, f = 1, and since mass is conserved, $\sigma = 0$. The mass current is $\rho\vec{v}$ so, for a one-component system

$$\frac{\partial\rho}{\partial t} + \vec{\nabla} \cdot (\rho\vec{v}) = 0 \quad . \qquad (2.6)$$

More generally, if there are N different components, γ, in the system with densities ρ_γ, then they satisfy the following equation:

$$\frac{\partial \rho_\gamma}{\partial t} + \vec{\nabla} \cdot (\rho_\gamma \vec{v}_\gamma) = \sum_{\delta=1}^{r} \nu_{\gamma\delta} J_\delta \qquad (2.7)$$

where: $\nu_{\gamma\delta} J_\delta$ is the production of component γ in the δth chemical reaction;

J_δ is the reaction rate for reaction δ;

$\nu_{\delta\gamma}$ is proportional to the stoichiometric coefficient with which component γ appears in the chemical equation for reaction δ .

Defining the following quantities:

$\rho \equiv \sum_{\gamma=1}^{N} \rho_\gamma$, the total density

$\vec{v} \equiv \sum_{\gamma=1}^{N} \frac{\rho_\gamma \vec{v}_\gamma}{\rho}$, the center of mass velocity

$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \vec{v} \cdot \vec{\nabla}$, the hydrodynamic (barycentric, material, or substantial) derivative

$\vec{J}_\gamma^{\,diff} \equiv \rho_\gamma (\vec{v}_\gamma - \vec{v})$, the diffusion flows,

the continuity equations become

$$\frac{D\rho_\gamma}{Dt} + \rho_\gamma \vec{\nabla} \cdot \vec{v} + \vec{\nabla} \cdot \vec{J}_\gamma^{\,diff} = \sum_{\delta=1}^{r} \nu_{\gamma\delta} J_\delta \qquad (2.8)$$

$$\frac{D\rho}{Dt} + \rho \vec{\nabla} \cdot \vec{v} = 0 \quad . \qquad (2.9)$$

For balance equations for variables other than mass,

$$\frac{\partial(f\rho)}{\partial t} = \rho \frac{\partial f}{\partial t} - f \vec{\nabla} \cdot (\rho \vec{v})$$

$$= \rho \frac{\partial f}{\partial t} + \rho\vec{v} \cdot \vec{\nabla} f - \vec{\nabla} \cdot (\rho f \vec{v})$$

$$= \sigma_F - \vec{\nabla} \cdot \vec{J}_F \,, \qquad (2.10)$$

so

$$\rho \frac{Df}{Dt} = \sigma_F - \vec{\nabla} \cdot (\vec{J}_F - \rho f \vec{v})$$

$$= \frac{\partial(f\rho)}{\partial t} + \vec{\nabla} \cdot (f\rho\vec{v}) \quad . \qquad (2.11)$$

For momentum balance, the source terms are known from continuum mechanics. They arise from external potentials and from internal forces which appear as a pressure. Using the material derivative, since the equation of motion applies to a volume of material, not of space:

$$\rho \frac{D\vec{v}}{Dt} = \sum_{\gamma=1}^{N} \rho_\gamma \vec{F}_\gamma - \vec{\nabla} \cdot \overleftrightarrow{P} \qquad (2.12)$$

where $\vec{F}_\gamma$ is the external force acting on component γ and $\overleftrightarrow{P}$ is the pressure tensor due to short range internal forces. Eqn. (2.12) can be written in the form of a balance equation as

$$\frac{\partial(\rho\vec{v})}{\partial t} + \vec{\nabla} \cdot (\overleftrightarrow{P} + \rho\vec{v}\vec{v}) = \sum_{\gamma=1}^{N} \rho_\gamma \vec{F}_\gamma \qquad (2.13)$$

where $\vec{v}\vec{v}$ is a tensor product (dyadic).

The term $\vec{\nabla} \cdot \overleftrightarrow{P}$ describes the transport of momentum out of a volume due to conduction processes (momentum transport through internal forces) while the term $\vec{\nabla} \cdot (\rho\vec{v}\vec{v})$ describes "convection" of momentum out of the volume due to the motion of the fluid. As in the equation of continuity, the right hand side is a source term. If there are no external forces ($\vec{F}_\gamma = 0$) then, from conservation of momentum,

$$\frac{\partial(\rho\vec{v})}{\partial t} + \vec{\nabla} \cdot (\overleftrightarrow{P} + \rho\vec{v}\vec{v}) = 0 \quad . \tag{2.14}$$

In the center of mass (Lagrange) frame, there is no convection term present and the pressure term gives the total momentum flow.

In general,

$$\overleftrightarrow{P} = \overleftrightarrow{P}^{e} + \overleftrightarrow{P}^{diss} \tag{2.15}$$

where $\overleftrightarrow{P}^{e}$ is the elastic part of the pressure tensor and includes a term $p\overleftrightarrow{U}$ ($\overleftrightarrow{U}$ is the unit tensor, $U_{ij} = \delta_{ij}$) which is simply the hydrostatic pressure and where $\overleftrightarrow{P}^{diss}$ is the dissipative contribution, related to viscosity. For fluids, $\overleftrightarrow{P}^{e} = p\overleftrightarrow{U}$ and at equilibrium only the elastic term is present (i.e., $\overleftrightarrow{P} = \overleftrightarrow{P}^{e}$).

For energy balance, eqn. (2.13) is multiplied by $\vec{v}$ to yield

$$\begin{aligned} \rho\vec{v} \cdot \frac{D\vec{v}}{Dt} = \rho\,\frac{D(\frac{1}{2}\,v^2)}{Dt} &= \vec{v} \cdot (\vec{\nabla} \cdot \overleftrightarrow{P}) + \sum_{\gamma=1}^{N} \rho_\gamma\,\vec{F}_\gamma \cdot \vec{v} \\ &= -\,\vec{\nabla} \cdot (\overleftrightarrow{P} \cdot \vec{v}) + \overleftrightarrow{P} : (\vec{\nabla}\vec{v}) + \sum_{\gamma=1}^{N} \rho_\gamma\,\vec{F}_\gamma \cdot \vec{v} \end{aligned} \tag{2.16}$$

where

$$\overleftrightarrow{P} : (\vec{\nabla}\vec{v}) \equiv \sum_{\alpha,\beta=1}^{3} P_{\alpha\beta} \frac{\partial v_\alpha}{\partial x_\beta} \quad . \tag{2.17}$$

As a local balance equation, this becomes

$$\frac{\partial(\frac{1}{2}\rho v^2)}{\partial t} + \vec{\nabla} \cdot (\frac{1}{2}\rho v^2 \vec{v} + \overleftrightarrow{P} \cdot \vec{v})$$

$$= \overleftrightarrow{P} : \vec{\nabla}\vec{v} + \sum_{\gamma=1}^{N} \rho_\gamma \vec{F}_\gamma \cdot \vec{v} \tag{2.18}$$

where $\frac{1}{2}\rho v^2 \vec{v}$ is a kinetic energy convection term and $\overleftrightarrow{P} \cdot \vec{v}$ is a kinetic energy conduction term.

The sources of kinetic energy are the external forces $\sum_{\gamma=1}^{N} \rho_\gamma \vec{F}_\gamma \cdot \vec{v}$ and the pressure tensor term $\overleftrightarrow{P} : \vec{\nabla}\vec{v}$ which represents the mechanical energy of compression.

For potential energy balance, if the external forces are due to some external potential, then

$$\vec{F}_\gamma = - \vec{\nabla}\psi_\gamma \tag{2.19}$$

where $\frac{\partial \psi_\gamma}{\partial t} = 0$ for time-independent forces.

Using eqn. (2.7) and the definition $\vec{J}_\gamma^{\,diff} = \rho_\gamma(\vec{v}_\gamma - \vec{v})$ where $\vec{v}$ is the barycentric velocity, if the total potential energy is written as

$$\rho\psi = \sum_{\gamma=1}^{N} \rho_\gamma \psi_\gamma \quad , \tag{2.20}$$

then

$$\frac{\partial(\rho\psi)}{\partial t} + \vec{\nabla}\cdot\left(\rho\psi\vec{v} + \sum_{\gamma=1}^{N} \psi_\gamma \vec{j}_\gamma^{\,diff}\right)$$
$$= -\sum_{\gamma=1}^{N} \rho_\gamma \vec{F}_\gamma \cdot \vec{v} - \sum_{\gamma=1}^{N} \vec{j}_\gamma^{\,diff}\cdot\vec{F}_\gamma + \sum_{\gamma=1}^{N}\sum_{\delta=1}^{r} \psi_\gamma \nu_{\gamma\delta} J_\delta \qquad (2.21)$$

where: $\rho\psi\vec{v}$ represents convection of potential energy;

$-\sum_{\gamma=1}^{N} \rho_\gamma \vec{F}_\gamma \cdot \vec{v}$ represents a sink due to conversion of potential energy to kinetic energy (an equal but opposite term appears in the kinetic energy equation);

$\sum_{\gamma=1}^{N} \psi_\gamma \vec{j}_\gamma^{\,diff}$ represents the potential energy of diffusion;

$-\sum_{\gamma=1}^{N} \vec{j}_\gamma^{\,diff}\cdot\vec{F}_\gamma$ represents the conversion of potential energy to internal energy by diffusion;

$\sum_{\gamma=1}^{N}\sum_{\delta=1}^{r} \psi_\gamma \nu_{\gamma\delta} J_\delta$ is a source due to change in potential energy due to chemical reactions.

In most cases, the last term will be zero, since the property of the particles responsible for the potential interaction usually remains unchanged in a chemical reaction (e.g., mass in a gravitational potential). In this case, the equation for total mechanical energy density $(\rho\psi + \frac{1}{2}\rho v^2)$ becomes

$$\frac{\partial(\rho\psi + \frac{1}{2}\rho v^2)}{\partial t} + \vec{\nabla}\cdot\left\{(\rho\psi + \frac{1}{2}\rho v^2)\vec{v} + \overleftrightarrow{P}\cdot\vec{v} + \sum_{\gamma=1}^{N} \psi_\gamma \vec{j}_\gamma^{\,diff}\right\}$$
$$= \overleftrightarrow{P} : \vec{\nabla}\vec{v} - \sum_{\gamma=1}^{N} \vec{j}_\gamma^{\,diff}\cdot\vec{F}_\gamma \quad . \qquad (2.22)$$

The presence of two source terms indicates the fact that the internal energy must be included in order to have energy conservation. In this equation, $\overleftrightarrow{P} : \vec{\nabla}\vec{v}$ is a source term arising from the conversion of kinetic energy to internal energy through compression and viscosity and $-\sum_{\gamma=1}^{N} \vec{j}_\gamma^{\,diff} \cdot \vec{F}_\gamma$ is a source due to conversion of internal energy to potential energy through diffusion processes.

The total energy e may be written as

$$\rho e = \frac{1}{2}\rho v^2 + \rho\psi + \rho u \qquad (2.23)$$

where the internal energy density u includes the energies of thermal agitation and of short range molecular interactions.

Due to energy conservation,

$$\frac{\partial(\rho e)}{\partial t} + \vec{\nabla} \cdot \vec{J}_e = 0 \quad . \qquad (2.24)$$

This implies that the source term for u is

$$\sigma_u = -\overleftrightarrow{P} : \vec{\nabla}\vec{v} + \sum_{\gamma=1}^{N} \vec{j}_\gamma^{\,diff} \cdot \vec{F}_\gamma \quad . \qquad (2.25)$$

In analogy to the current for mechanical energy, the current for total energy has a convective term $\rho e\vec{v}$, the previously discussed mechanical and potential flux terms, $\overleftrightarrow{P} \cdot \vec{v}$ and $\sum_{\gamma=1}^{N} \psi_\gamma \vec{j}_\gamma^{\,diff}$, and a new internal energy flux $\vec{J}_q$. This defines $\vec{J}_q$ which is commonly called the heat flux.

Thus, the balance equation for total energy is

$$\frac{\partial(\rho e)}{\partial t} + \vec{\nabla} \cdot \left(\rho e\vec{v} + \overleftrightarrow{P} \cdot \vec{v} + \sum_{\gamma=1}^{N} \psi_\gamma \vec{J}_\gamma^{\,diff} + \vec{J}_q\right) = 0 \tag{2.26}$$

and the balance equation for internal energy is

$$\frac{\partial(\rho u)}{\partial t} + \vec{\nabla} \cdot (\rho u\vec{v} + \vec{J}_q) = - \overleftrightarrow{P} : \vec{\nabla}\vec{v} + \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{F}_\gamma \quad . \tag{2.27}$$

This is simply the first law of thermodynamics. It can be put in more familiar form as

$$\frac{Du}{Dt} = \frac{Dq}{Dt} - p\,\frac{Dv}{Dt} - v\,\overleftrightarrow{\Pi} : \vec{\nabla}\vec{v} + v \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{F}_\gamma \tag{2.28}$$

where: $v \equiv \rho^{-1}$ is the specific volume;

p is the scalar hydrostatic pressure;

$\overleftrightarrow{\Pi} \equiv \overleftrightarrow{P} - p\overleftrightarrow{U}$ is the pressure tensor without the hydrostatic part;

q is the heat per unit mass, defined by

$\rho\,\frac{Dq}{Dt} + \vec{\nabla} \cdot \vec{J}_q = 0$.

The final mechanical density which will be considered is angular momentum density. If the angular momentum density is $\vec{J}$, then this can also be expressed in terms of an angular momentum density tensor $\overleftrightarrow{J}$ with components

$$\overleftrightarrow{J}_{\alpha\beta} = \sum_{\gamma=1}^{3} \varepsilon_{\alpha\beta\gamma}\, \vec{J}_\gamma \tag{2.29}$$

where $\varepsilon_{\alpha\beta\gamma}$ is the antisymmetric Levi-Civita tensor.

The local form of the angular momentum conservation equation is, assuming no external forces,

$$\rho \frac{D\overleftrightarrow{J}_{\alpha\beta}}{Dt} = - \sum_{\gamma=1}^{3} \frac{\partial}{\partial r_\gamma} (r_\alpha \overleftrightarrow{P}_{\gamma\beta} - r_\beta \overleftrightarrow{P}_{\gamma\alpha}) \quad . \qquad (2.30)$$

The right hand side represents the flow of angular momentum due to the torque exerted on a mass element by the pressure tensor.

The angular momentum density tensor $\overleftrightarrow{J}$ is now split into two parts

$$\overleftrightarrow{J} = \overleftrightarrow{L} + \overleftrightarrow{S} \qquad (2.31)$$

where $L_{\alpha\beta}(\equiv r_\alpha v_\beta - r_\beta v_\alpha)$ is the usual angular momentum density due to the circulation of the material. In classical hydrodynamics, the material has no micro-structure and $\overleftrightarrow{L}$ is the total angular momentum. More generally, there is another contribution $\overleftrightarrow{S}_{\alpha\beta}$ which is the internal angular momentum of the material. It arises from the fact that the molecules making up the material can have angular momentum (spin) without having a macroscopic fluid velocity. $\overleftrightarrow{S}_{\alpha\beta}$ can be written as

$$\overleftrightarrow{S}_{\alpha\beta} = \Theta \overleftrightarrow{\omega}_{\alpha\beta} \qquad (2.32)$$

where $\overleftrightarrow{\omega}_{\alpha\beta}$ is the antisymmetric rotation tensor corresponding to $\vec{\omega}$, the angular velocity, and where Θ is the average moment of inertia per unit mass.

Taking the equation of motion

$$\frac{D\vec{v}}{Dt} = -\vec{\nabla} \cdot \overleftrightarrow{P} \tag{2.33}$$

and multiplying by $\vec{r}$ and subtracting a transposed term gives

$$r_\alpha \frac{Dv_\beta}{Dt} - r_\beta \frac{Dv_\alpha}{Dt} = \frac{D\overleftrightarrow{L}_{\alpha\beta}}{Dt} \quad . \tag{2.34}$$

Therefore,

$$\frac{D\overleftrightarrow{L}_{\alpha\beta}}{Dt} = \{-\sum_{\gamma=1}^{3} \frac{\partial}{\partial r_\gamma} (r_\alpha \overleftrightarrow{P}_{\gamma\beta} - r_\beta \overleftrightarrow{P}_{\gamma\alpha})\} + (\overleftrightarrow{P}_{\alpha\beta} - \overleftrightarrow{P}_{\beta\alpha}) \quad . \tag{2.35}$$

Hence, from angular momentum conservation,

$$\frac{D\overleftrightarrow{S}}{Dt} = -2\, \overleftrightarrow{P}^{(a)} \tag{2.36}$$

where $\overleftrightarrow{P}^{(a)}$ is the antisymmetric part of $\overleftrightarrow{P}$.

If $\overleftrightarrow{S} = 0$, the constituent particles have no angular momentum and $\overleftrightarrow{P}_{\alpha\beta} = \overleftrightarrow{P}_{\beta\alpha}$, as is usually assumed in continuum mechanics. The antisymmetric part of the pressure tensor is due to the internal body torques which are proportional to volume. These can arise if the constituent molecules are not spherically symmetric. If the molecules are spherical or if they can be regarded as structureless (as in a dilute gas), then $\overleftrightarrow{P}_{\alpha\beta} = \overleftrightarrow{P}_{\beta\alpha}$ and $\overleftrightarrow{S}_{\alpha\beta}$ = constant. This means that internal angular momentum $\overleftrightarrow{S}$ and external angular momentum are separately conserved. In viscous fluids, however, the two angular momenta would not be expected to be independent. Instead, viscous effects would result in transfer between internal and external angular momentum.

To investigate the effect of $\overleftrightarrow{P}^{(a)}_{\alpha\beta}$ on the motion of the fluid, it is assumed that the pressure tensor is a generalization of the tensor for a Newtonian fluid. That is, the pressure tensor components are linear functions of the appropriate velocity gradients. For example, if the viscous part of $\overleftrightarrow{P}$ is separated and elasticity is neglected,

$$\overleftrightarrow{P} = p\overleftrightarrow{U} + \overleftrightarrow{\Pi}$$

where $\overleftrightarrow{\Pi}$ is the viscous pressure tensor, written as

$$\overleftrightarrow{\Pi} = \Pi_o \overleftrightarrow{U} + \overleftrightarrow{\Pi}^{(s)}_{tr=0} + \overleftrightarrow{\Pi}^{(a)} \quad . \qquad (2.37)$$

Then the assumption of linear constitutive equations is:

$$\Pi_o = - \eta_{vol} (\vec{\nabla} \cdot \vec{v})$$

$$\overleftrightarrow{\Pi}^{(s)}_{tr=0} = - 2\, \eta_{shear} (\vec{\nabla}\vec{v})^{(s)}_{tr=0}$$

$$\{\text{i.e., } (\overleftrightarrow{\Pi}^{(s)}_{tr=0})_{\alpha\beta} = - \eta_{shear}\left(\frac{\partial v_\alpha}{\partial r_\beta} + \frac{\partial v_\beta}{\partial r_\alpha} - 2\vec{\nabla} \cdot \vec{v}\, \delta_{\alpha\beta}\right)\}$$

$$\overleftrightarrow{\Pi}^{(a)} = - \eta_{rot} (\vec{\nabla} \times \vec{v} - 2\overleftrightarrow{\omega}) \qquad (2.38)$$

($\vec{\nabla} \times \vec{v}$ is regarded as an antisymmetric tensor)

where three phenomenological coefficients have been introduced:

η_{vol}, the bulk viscosity

η_{shear}, the shear viscosity

η_{rot}, the rotational viscosity.

It is the substitution of $-\eta_{rot}(\vec{\nabla} \times \vec{v} - 2\overleftrightarrow{\omega})$ for $\overleftrightarrow{\Pi}^{(a)}$ that differs from the usual assumption that $\overleftrightarrow{\Pi}^{(a)} = 0$.

If it is assumed that the viscosity coefficients are independent of position, then the modified Navier-Stokes equation is

$$\rho \frac{D\vec{v}}{Dt} = -\vec{\nabla}p + \eta_{shear}\nabla^2\vec{v} + \left(\frac{1}{3}\eta_{shear} + \eta_{vol}\right)\vec{\nabla}(\vec{\nabla}\cdot\vec{v}) + \eta_{rot}\vec{\nabla}\times(2\vec{\omega} - \vec{\nabla}\times\vec{v}) \quad . \qquad (2.39)$$

In order to see the physical significance of the new term $\eta_{rot}\vec{\nabla}\times(2\vec{\omega} - \vec{\nabla}\times\vec{v})$, two special cases are considered.

First, pure expansion (no shears or rotations of the fluid) is considered. Then,

$$\vec{v} = \alpha\vec{r}$$

$$(\vec{\nabla}\vec{v})^{(s)}_{tr=0} = 0$$

$$\vec{\nabla}\times\vec{v} = 0$$

$$\Pi_o = -\eta_{vol}(\vec{\nabla}\cdot\vec{v}) = -\eta_{vol}(3\alpha) \neq 0 \quad . \qquad (2.40)$$

Thus, volume viscosity appears in the equation of motion but shear viscosity does not. If the molecules have spin initially, the rotational viscosity will transmit the angular momentum and cause $\vec{\nabla} \times \vec{v}$ to become non-zero (i.e., cause the fluid motion to depart from pure expansion).

Secondly, rigid body rotation, where $\vec{v} = \vec{b} \times \vec{r}$ ($\vec{b}$ a constant vector) is considered . Then $(\vec{\nabla}\vec{v})^{(s)}_{tr=0} = 0$, $\vec{\nabla} \cdot \vec{v} = 0$, and the only non-zero term in the viscous pressure tensor is $\overleftrightarrow{\Pi}^{(a)} = - \eta_{rot}(\vec{\nabla} \times \vec{v} - 2\overleftrightarrow{\omega})$. The equation of motion for $\overleftrightarrow{S}$ is

$$\rho \frac{D\overleftrightarrow{S}}{Dt} = -2 \overleftrightarrow{\Pi}^{(a)} = 2\eta_{rot} (\vec{\nabla} \times \vec{v} - 2\overleftrightarrow{\omega}) \quad . \quad (2.41)$$

Considering a case where the rotation of the fluid is constant ($\vec{\nabla} \times \vec{v} = 2\vec{b}$ where $\frac{\partial \vec{b}}{\partial t} = 0$), then since $\vec{S} = \theta\vec{\omega}$,

$$\frac{d\vec{\omega}}{dt} = - \frac{4\eta_{rot}}{\rho\theta} (\vec{\omega} - \vec{b}) \quad . \quad (2.42)$$

If initially $\vec{\omega} = 0$ (the molecules have no initial spin) then

$$\vec{\omega} = \vec{b}(1 - e^{-t/\tau}) \quad (2.43)$$

where $\tau = \dfrac{\rho\theta}{4\eta_{rot}}$.

Thus, after a time $t \gg \tau$, $2\vec{\omega}$ is essentially equal to $\vec{\nabla} \times \vec{v}$ and the rotational viscous effects become negligible and the internal and translational angular momenta are in equilibrium.

§2.2 Models of Real Materials

In many applications, it is possible to consider either elastic or viscous effects separately. For example, in hydrodynamics, it is assumed that the fluid has no elasticity. Also, in solid mechanics it is often assumed that there can be no flow or permanent deformation and hence the stress tensor is purely elastic. However, there are situations where both viscous and elastic phenomena are important. Examples are plastic substances which flow under sustained stress yet which retain many of the elastic properties of other solids. There are also some liquids with appreciable elastic properties. For example, there are certain "gels" which combine fluidity (low shear viscosity) with elasticity in such a way that immersing a pendulum in the gel actually decreases its period of oscillation (cf. Frenkel [5] p.218).

In purely elastic deformations, the quantities of interest are the linear strain tensor

$$\overleftrightarrow{\varepsilon}_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial u_\alpha(\vec{r},t)}{\partial r_\beta} + \frac{\partial u_\beta(\vec{r},t)}{\partial r_\alpha}\right) \quad , \qquad (2.44)$$

where $u_\alpha(\vec{r},t)$ is the α component of the vector $\vec{u}$ connecting the equilibrium position $\vec{r}$ of a material with its position after deformation, and the stress tensor

$$\overleftrightarrow{\sigma} = -\overleftrightarrow{P} = -p\overleftrightarrow{U} - \overleftrightarrow{\Pi} \quad . \qquad (2.45)$$

In a linear elasticity theory, the material obeys a generalized Hooke's Law:

$$\overleftrightarrow{\sigma}_{\alpha\beta} = \overleftrightarrow{C}_{\alpha\beta\gamma\delta}\, \overleftrightarrow{\varepsilon}_{\gamma\delta}$$

where the $C_{\alpha\beta\gamma\delta}$ are elastic constants.

In viscoelastic theory, the rate of strain tensor

$$\overleftrightarrow{\Lambda}_{\alpha\beta} = \frac{\partial \overleftrightarrow{\varepsilon}_{\alpha\beta}}{\partial t} = \frac{1}{2}\left(\frac{\partial v_\alpha}{\partial r_\beta} + \frac{\partial v_\beta}{\partial r_\alpha}\right) \qquad (2.46)$$

must also be considered. Separating $\overleftrightarrow{\Lambda}$ into an elastic part $\overleftrightarrow{\Lambda}^{(e)}$ and a viscous part $\overleftrightarrow{\Lambda}^{(v)}$ gives $\overleftrightarrow{\Lambda} = \overleftrightarrow{\Lambda}^{(e)} + \overleftrightarrow{\Lambda}^{(v)}$.

The equation relating the viscous part of the rate of strain to the stress tensor $\overleftrightarrow{\sigma}$ is

$$\overleftrightarrow{\sigma} = - p\overleftrightarrow{U} + 2\eta_{shear}\, \overleftrightarrow{\Lambda}^{(v)} + (\eta_{vol} - \tfrac{2}{3}\eta_{shear})\, \mathrm{Tr}(\overleftrightarrow{\Lambda}^{(v)})\overleftrightarrow{U} \qquad (2.47)$$

where Tr denotes trace. This is for a Newtonian fluid where the stress tensor is a linear function of the rate of strain tensor. This can be written as

$$\overleftrightarrow{\sigma} + p\overleftrightarrow{U} = - \overleftrightarrow{\Pi} = 2\eta_{shear}\left[\overleftrightarrow{\Lambda}^{(v)} - \frac{1}{3}\mathrm{Tr}(\overleftrightarrow{\Lambda}^{(v)})\overleftrightarrow{U}\right] + \eta_{vol}\, \mathrm{Tr}(\overleftrightarrow{\Lambda}^{(v)})\overleftrightarrow{U} . \qquad (2.48)$$

The apparently arbitrary regrouping of viscous terms is to separate viscous effects of changes in shape (first term) and changes in volume (second term).

If the fluid is isotropic, the equation for linear viscosity is

$$\overleftrightarrow{\sigma}_{\alpha\beta} + p\,\overleftrightarrow{U}_{\alpha\beta} = G\left(\frac{\partial u_\alpha}{\partial r_\beta} + \frac{\partial u_\beta}{\partial r_\alpha} - \frac{2}{3}(\vec{\nabla}\cdot\vec{u})\delta_{\alpha\beta}\right) + $$

$$+ B(\vec{\nabla}\cdot\vec{u})\delta_{\alpha\beta} \qquad (2.49)$$

where, as before, the separation is into shape and volume dependent terms. Here, G is the shear modulus and B is the bulk modulus.

Taking the time derivative of the elastic equation and using the definition $\frac{\partial\vec{u}}{\partial t} = \vec{v}^{(e)}$,

$$\frac{1}{G}\frac{\partial}{\partial t}(\overleftrightarrow{\sigma} + p\,\overleftrightarrow{U}) = 2[\overleftrightarrow{\Lambda}^{(e)} - \frac{1}{3}\mathrm{Tr}(\overleftrightarrow{\Lambda}^{(e)})\overleftrightarrow{U}]$$

$$+ \frac{B}{G}(\vec{\nabla}\cdot\vec{v}^{(e)})\overleftrightarrow{U} \quad . \qquad (2.50)$$

Also, since $\mathrm{Tr}(\overleftrightarrow{\Lambda}^{(v)}) = \vec{\nabla}\cdot\vec{v}^{(v)}$ may be taken as a definition of $\vec{v}^{(v)}$, and with $\vec{v} = \vec{v}^{(e)} + \vec{v}^{(v)}$, the viscosity and elasticity equations can be combined to yield

$$\left\{\frac{1}{\eta_{shear}} + \frac{1}{G}\frac{\partial}{\partial t}\right\}\{\overleftrightarrow{\sigma} + p\,\overleftrightarrow{U}\}$$

$$= 2[\overleftrightarrow{\Lambda} - \frac{1}{3}\mathrm{Tr}(\overleftrightarrow{\Lambda})\overleftrightarrow{U}] + A(\vec{\nabla}\cdot\vec{v})\overleftrightarrow{U} \quad . \qquad (2.51)$$

This equation, valid for an isotropic, Newtonian viscoelastic fluid, also assumes an equal relaxation time approximation, $A = \frac{\eta_{vol}}{\eta_{shear}} = \frac{B}{G}$, which is justified in certain simple models (cf. Lodge [6] p. 110).

It should be emphasized that eqn (2.51) is valid only as one possible model and is not a general equation for any effect involving viscous and elastic properties. The essential assumption (other than that of linear constitutive equations) was that the effects of viscous and elastic properties can be considered separately and that the net effect is simply additive. This corresponds to the "Maxwell model" for viscoelastic solids. This model is most simply illustrated by the one-dimensional apparatus in Fig. 2.1. The spring (elastic effect) and the dashpot (viscous effect) are assumed linear as

$\sigma = G\varepsilon$ for the spring alone

and $\sigma = \eta \frac{\partial\varepsilon}{\partial t}$ for the dashpot alone,

where ε is the displacement. The equation of motion for the system is

$$\{\frac{1}{\eta} + \frac{1}{G}\frac{\partial}{\partial t}\}\sigma = \frac{\partial\varepsilon}{\partial t} \qquad (2.52)$$

in obvious analogy to the tensor equation for an isotropic viscoelastic fluid (eqn. (2.51)).

However, it is obvious that there are other viscoelastic models possible and, in fact, the Maxwell model is inadequate to explain viscoelastic effects in real solids. For example, another model is the Kelvin model, illustrated in Fig. 2.2. The equation of motion for this

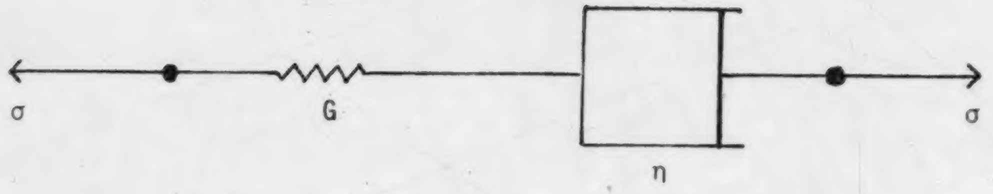


Fig. 2.1 One-dimensional Maxwell model for viscoelasticity

σ = stress

G = spring constant

η = dashpot viscosity

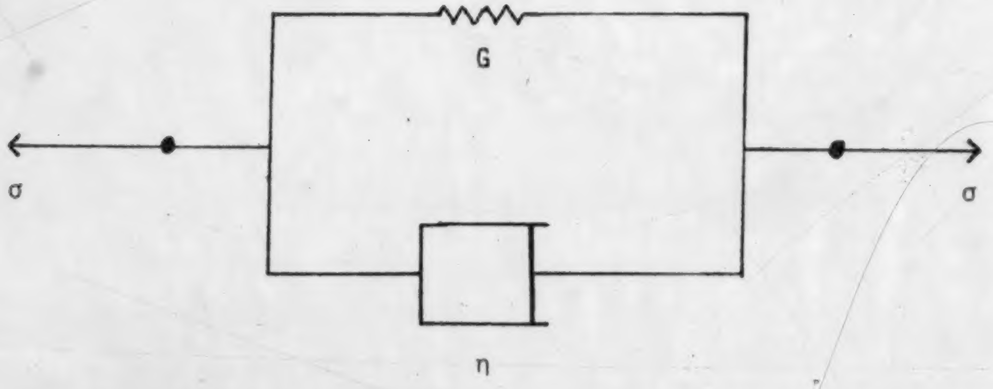


Fig. 2.2 One-dimensional Kelvin model for viscoelasticity

σ = stress

G = spring constant

η = dashpot viscosity

model is, in one dimension,

$$\sigma = G\varepsilon + \eta \frac{\partial \varepsilon}{\partial t} \quad . \qquad (2.53)$$

§2.3 Entropy Balance and the Second Law of Thermodynamics

After considering the balance equations for the various mechanical quantities, the connection to thermodynamics is made by writing an equation for entropy balance. For a macroscopic thermodynamic system, the entropy S which has the following properties, is introduced.

Firstly, it is additive (the entropy of the sum of two systems is the sum of their entropies). This implies that a differential change in entropy dS can be written as

$$dS = dS_e + dS_i \quad . \qquad (2.54)$$

Here, dS_e arises from the exchange of entropy with the system's surroundings and dS_i comes from internal production or destruction of entropy.

Secondly, dS_e can be positive, negative, or zero depending upon the system's interaction with its surroundings, but $dS_i \geq 0$ always. Also, $dS_i = 0$ only for reversible changes. Thus, for an isolated system, $dS_e = 0$ since there is no interaction with the surroundings, and hence $dS \geq 0$ (second law of thermodynamics).

Finally, if the system is free to receive heat from a reservoir at temperature T, but is otherwise isolated (closed system), then $dS_e = \frac{dQ}{T}$. Hence, $dS \geq \frac{dQ}{T}$, which is an alternate form of the second law.

Under the assumption of local equilibrium, it is possible to introduce a local entropy density s, an internal entropy source density σ_s, and an entropy flow density $\vec{J}_s$ which includes both internal and external entropy flows. Then,

$$S = \int_V \rho s \, dV$$

$$\frac{dS_e}{dt} = - \int_\Sigma \vec{J}_s \cdot d\vec{\Sigma}$$

$$\frac{dS_i}{dt} = \int_V \sigma_s \, dV \quad . \tag{2.55}$$

Since, in local equilibrium situations $s = s(u,v,c_\gamma)$ is a continuous function of the internal energy density u, the specific volume $v \equiv \rho^{-1}$, and the mass fraction $c_\gamma \equiv \rho_\gamma/\rho$, Gauss' theorem can be applied to yield

$$\frac{dS}{dt} = \int_V \frac{\partial(\rho s)}{\partial t} dV = \frac{dS_e}{dt} + \frac{dS_i}{dt}$$

$$= - \int_V (\vec{\nabla} \cdot \vec{J}_s) dV + \int_V \sigma_s \, dV \quad . \tag{2.56}$$

Hence, in local form,

$$\frac{\partial(\rho s)}{\partial t} + \vec{\nabla} \cdot \vec{J}_s = \sigma_s \qquad (2.57)$$

where $\sigma_s \geq 0$. Alternatively,

$$\rho \frac{Ds}{Dt} + \vec{\nabla} \cdot (\vec{J}_s - \rho s \vec{v}) = \sigma_s \quad . \qquad (2.58)$$

To find explicit forms of $\vec{J}_s$ and σ_s, the Gibbs-Duhem formula must be used. Using

$$ds(u,v,c_\gamma) = \frac{\partial s}{\partial u} du + \frac{\partial s}{\partial v} dv + \sum_{\gamma=1}^{N} \frac{\partial s}{\partial c_\gamma} dc_\gamma \qquad (2.59)$$

and the definitions

$$\left(\frac{\partial s}{\partial u}\right)_{v,c_\gamma} = T^{-1} \; ; \; \left(\frac{\partial s}{\partial v}\right)_{u,c_\gamma} = pT^{-1} \; ; \; \left(\frac{\partial s}{\partial c_\gamma}\right)_{v,u} = -\mu_\gamma T^{-1}$$

where p is the equilibrium pressure and μ_γ is the chemical potential per unit mass for component γ, the Gibbs-Duhem relation is found as

$$Tds = du + pdv - \sum_{\gamma=1}^{N} \mu_\gamma \, dc_\gamma \quad . \qquad (2.60)$$

This, by virtue of the local equilibrium condition, is valid near complete equilibrium. In particular, looking at a volume element in the center of mass frame gives

$$T \frac{Ds}{Dt} = \frac{Du}{Dt} + p \frac{Dv}{Dt} - \sum_{\gamma=1}^{N} \mu_\gamma \frac{Dc_\gamma}{Dt} \quad . \qquad (2.61)$$

Substituting from eqns. (2.8) and (2.27) gives

$$\rho \frac{Ds}{Dt} = -\frac{1}{T} \vec{\nabla} \cdot \vec{J}_q - \frac{1}{T} \overleftrightarrow{\Pi} : \vec{\nabla}\vec{v} + \frac{1}{T} \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{F}_\gamma$$

$$+ \frac{1}{T} \sum_{\gamma=1}^{N} \mu_\gamma \vec{\nabla} \cdot \vec{J}_\gamma^{\,diff} - \frac{1}{T} \sum_{\delta=1}^{r} J_\delta A_\delta \qquad (2.62)$$

where $A_\delta = \sum_{\gamma=1}^{N} \nu_{\gamma\delta} \mu_\gamma$ is the chemical affinity. The terms in this equation can also be rearranged into a balance equation:

$$\frac{\partial(\rho s)}{\partial t} + \vec{\nabla} \cdot \vec{J}_s = \sigma_s \qquad (2.63)$$

where $$\vec{J}_s = \{\rho s \vec{v} + \frac{\vec{J}_q}{T} - \sum_{\gamma=1}^{N} \mu_\gamma \frac{\vec{J}_\gamma^{\,diff}}{T}\} \qquad (2.64)$$

and

$$\sigma_s = \vec{J}_q \cdot \vec{\nabla}(\frac{1}{T}) + \frac{1}{T} \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{F}_\gamma - \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{\nabla}(\frac{\mu_\gamma}{T})$$

$$- \frac{1}{T} \overleftrightarrow{\Pi} : \vec{\nabla}\vec{v} - \frac{1}{T} \sum_{\delta=1}^{r} J_\delta A_\delta \quad . \qquad (2.65)$$

The separation into sources and currents is determined by the conditions that σ_s must be zero if the system is in equilibrium and also that the separation must be invariant under a Galilei transformation, since the concept of irreversibility is independent of frame of reference. Note also that, since for a closed system

$$\frac{Ds}{Dt} = - \int_\Sigma \vec{J}_s \cdot d\vec{\Sigma} + \int_V \sigma_s \, dV$$

$$\geq \int_\Sigma \vec{J}_s \cdot d\vec{\Sigma} \quad , \qquad (2.66)$$

$\vec{v} = 0$, and $\vec{J}_\gamma^{\,diff} = 0$ at the surface,

$$\frac{Ds}{Dt} \geq \int_\Sigma \frac{\vec{J}_q \cdot d\vec{\Sigma}}{T} \quad \text{(Clausius Carnot theorem)} \quad . \qquad (2.67)$$

The structure of the entropy source term is that of a sum of products of thermodynamic fluxes and generalized thermodynamic forces. The various terms are:

(1) $\vec{J}_q \cdot \vec{\nabla}(\frac{1}{T})$, a heat conduction term. (The flux is the heat current $\vec{J}_q$ and the force is $\vec{\nabla}(\frac{1}{T})$ related to the temperature gradient.)

(2) $\frac{1}{T} \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{F}_\gamma$, a term describing diffusion in an external field. (The flux is $\vec{J}_\gamma^{\,diff}$ and the force is $\frac{1}{T} \vec{F}_\gamma$, the external force.)

(3) $- \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{\nabla}(\frac{\mu_\gamma}{T})$, a diffusion term due to chemical potentials. (The flux is again $\vec{J}_\gamma^{\,diff}$ but the force is $- \vec{\nabla}(\frac{\mu_\gamma}{T})$, related to the inhomogeneity of the chemical potential.)

(4) $- \frac{1}{T} \overleftrightarrow{\Pi} : \vec{\nabla}\vec{v}$, a viscous pressure term. (The flux is $\overleftrightarrow{\Pi}$, related to momentum flow, and the force is $- \frac{1}{T} \vec{\nabla}\vec{v}$, the velocity gradient.)

(5) $- \frac{1}{T} \sum_{\delta=1}^{r} J_\delta A_\delta$, a chemical reaction term. (The flux is J_δ, the reaction rate, and the force which drives the reaction is the affinity A_δ.)

Note that it is possible to put some of these terms in a different form. For example, by redefining the heat flow $\vec{j}_q$ to include heat transfer due to diffusion, the term involving $\vec{\nabla}T$ and $\vec{j}_\gamma^{\,diff}$ can be eliminated.

In general, the entropy source σ_s may be written as

$$\sigma_s = \sum_i Y_i \, X_i \tag{2.68}$$

where the Y_i are the thermodynamic fluxes ($\vec{j}_\gamma^{\,diff}$, $\vec{j}_q$, $\overleftrightarrow{\Pi}$, etc.) which are "driven" by the conjugate thermodynamic forces ($\vec{\nabla}(\frac{\mu_\gamma}{T})$, $\vec{F}_\gamma$, $\nabla(\frac{1}{T})$, $\frac{1}{T}\,\vec{\nabla}\vec{v}$, etc.).

The next chapter will include a study of the structure of the entropy source term.

CHAPTER III

LINEAR PHENOMENOLOGICAL LAWS

§3.1 Choosing "Forces" and "Fluxes"

From the balance equation for entropy, it is known that the entropy production σ_s can be written as a sum of products of generalized thermodynamic fluxes Y_i and forces X_i. For the conditions considered in Chapter II,

$$\sigma_s = - \vec{J}_q \cdot \left(\frac{1}{T^2} \vec{\nabla} T\right) + \frac{1}{T} \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{F}_\gamma - \sum_{\gamma=1}^{N} \vec{J}_\gamma^{\,diff} \cdot \vec{\nabla}\left(\frac{\mu_\gamma}{T}\right)$$

$$- \frac{1}{T} \overleftrightarrow{\Pi} : \vec{\nabla}\vec{v} - \frac{1}{T} \sum_{\delta=1}^{r} J_\delta A_\delta$$

$$= \sum_{i=1}^{5} X_i Y_i \quad . \qquad (3.1)$$

Intuitively, for each term in eqn. (3.1), it is possible to separate the force and the flux. For example, for thermal conduction, $X_1 Y_1 = - \vec{J}_q \cdot (\frac{1}{T^2} \vec{\nabla} T)$ and some multiple of $\vec{J}_q$ is interpreted as the flux Y_1, while a multiple of $\vec{\nabla} T$ is regarded as the force X_1. The choice of proportionality constants must be made in such a way that $X_1 Y_1 = - \frac{1}{T^2} \vec{J}_q \cdot \vec{\nabla} T$. For example, one choice would be $X_1 = - \vec{\nabla} T$ and $Y_1 = \frac{1}{T^2} \vec{J}_q$ and an alternative valid choice would be $X_1' = \vec{\nabla}(\frac{1}{T})$ and $Y_1' = \vec{J}_q$. Such choices are essentially ones of convenience and have no effect on the physical interpretation.

The intuitive separation of forces and fluxes seems to be based upon the concepts of cause and effect. In a given experimental situation, it seems possible to establish, as a constraint, one of the components of each term in σ_s. For example, temperature gradients are established by putting different boundaries of the system in contact with reservoirs at different temperatures, external fields are applied, gradients in chemical potentials are maintained by concentration differences, shear flow (gradients in $\vec{v}$) is maintained by suitable boundary conditions, or non-zero affinities are created in chemical systems by adding reactants. These externally applied forces are then seen as "causing" the corresponding fluxes. However, the clear-cut distinction between controlled external forces, such as $\vec{\nabla}T$, and the "resulting" fluxes, such as $\vec{J}_q$, disappears when it is realized that $\vec{\nabla}T$ itself is not controlled, but rather boundary conditions are established. For example, consider conduction of heat in a slab with its ends maintained at different temperatures. In order to find $T(\vec{r},t)$, and hence to find $\vec{\nabla}T$, it is necessary to solve a boundary value problem involving the specific heat of the material and the heat current $\vec{J}_q$, as well as the boundary conditions. Hence, the "applied force" $\vec{\nabla}T$ implicitly involves the "resultant flux" $\vec{J}_q$, a highly unsatisfactory state of affairs! It is obviously desirable to have a much more specific criterion for separating forces from fluxes in thermodynamics if these concepts are to have any real meaning.

The main reason for wanting a clear definition of force and flux is to judge the validity of certain symmetry relations in the phenomenological laws connecting the forces to the fluxes (Onsager reciprocal relations).

First, the form of these phenomenological laws will be stated. In equilibrium, the applied forces are zero and, in agreement with the usual concept of equilibrium, it is desirable that the resultant fluxes should also be zero. If it is assumed that each flux Y_i can depend upon all of the forces X_i, then

$$Y_i = Y_i(X_1,\ldots,X_n) \tag{3.2}$$

if there are n separate forces.

In equilibrium,

$$Y_i(0,0,\ldots,0) = 0 \quad . \tag{3.3}$$

Hence, for sufficiently small deviations from equilibrium, the form of the phenomenological relationship between forces and fluxes is

$$Y_i = \sum_{k=1}^{n} L_{ik} X_k \tag{3.4}$$

where the L_{ik} are constant phenomenological coefficients.

This general form includes the usual phenomenological linear equations of macroscopic physics as special cases. For example, Ohm's law ($\vec{J} = \sigma\vec{E}$ where σ is electrical conductivity), Fourier's law ($\vec{j}_q = -\lambda\vec{\nabla}T$ where λ is thermal conductivity), Fick's law of diffusion ($\vec{j}_\gamma^{\,diff} = -D\vec{\nabla}n_\gamma$ where n_γ is concentration of "γ"), and Newton's law of friction ($\sigma_{\alpha\beta} = \frac{\eta}{2}\left(\frac{\partial v_\alpha}{\partial r_\beta} + \frac{\partial v_\beta}{\partial r_\alpha}\right)$ where σ is stress tensor) are all special cases in which only one flux and one force are present. More generally, processes involving several forces at once may be considered. Thermoelectric, electrokinetic, and galvanomagnetic phenomena are examples of such complex processes, as well as thermal or electric conduction in cases where the current is in a different direction from the applied force.

The above examples generalize, for vector phenomena, to

$$
\begin{aligned}
\vec{J} &= \overleftrightarrow{\sigma} \cdot \vec{E} \\
\vec{j}_q &= \overleftrightarrow{\lambda} \cdot (\vec{\nabla}T) \\
\vec{j}_\gamma^{\,diff} &= -\overleftrightarrow{D} \cdot (\vec{\nabla}n_\gamma) \\
\overleftrightarrow{\sigma}_{\alpha\beta} &= \frac{\overleftrightarrow{\eta}_{\alpha\beta\gamma\delta}}{2}\left(\frac{\partial v_\alpha}{\partial r_\beta} + \frac{\partial v_\beta}{\partial r_\alpha}\right) \qquad (3.5)
\end{aligned}
$$

where the constant coefficients have become constant tensors.

For other examples, see the article by Miller [7].

The positive definite property of the entropy production results in a condition upon the coefficients L_{ik},

$$\sum_i X_i Y_i = \sum_{i,k} L_{ik} X_i X_k \geq 0 \quad . \tag{3.6}$$

In order for this bilinear form to be positive semi-definite, all the diagonal elements L_{ii} must be positive or zero. The restrictions on the off diagonal elements are less severe and are of the form $L_{ii} L_{kk} \geq \frac{1}{4} (L_{ik} + L_{ki})^2$. Hence, the off-diagonal elements $L_{ik} (i \neq k)$ are not necessarily positive or zero. For a derivation of these conditions see, for example, Hoffman and Kunze [8].

Note that it is only the symmetric part of the matrix $\overleftrightarrow{L}$ which contributes to the entropy production. That is, if $\overleftrightarrow{L} = \overleftrightarrow{L}^{(a)} + \overleftrightarrow{L}^{(s)}$, where $L^{(s)}$ is the symmetric part of $\overleftrightarrow{L}$ and $\overleftrightarrow{L}^{(a)}$ the antisymmetric part, then

$$\sigma_s = \sum_{i,k} L_{ik} X_i X_k = \sum_{i,k} L_{ik}^{(s)} X_i X_k \tag{3.7}$$

and there is freedom to add an antisymmetric part to $\overleftrightarrow{L}$ without changing σ_s. This is important when considering the dependence of the Onsager relations upon the particular choice of forces and fluxes.

Also in most systems, the components of the various fluxes do not depend on all of the forces. This is a consequence of what is called Curie's principle. Loosely stated, Curie's principle asserts that, in an isotropic medium, fluxes and forces of different tensorial character do not couple. Some care must be taken in interpreting this principle (cf. Truesdell [9] pp. 134-137; de Groot and Mazur [4] Chapter VI) since it is not immediately obvious what is meant by tensorial character. For example, a second rank tensor may be regarded as being composed of a scalar (the trace of the tensor), a vector (the antisymmetric part of the tensor, an axial vector), and a symmetric tensor. However, with proper care, this symmetry of an isotropic system can be used to reduce the number of independent relationships between the forces and fluxes. For example, in an isotropic fluid, heat conduction and diffusion are vector phenomena, and viscous phenomena or chemical reactions are tensor and scalar phenomena respectively.

In systems which are not isotropic, it is also possible to use the symmetries of the system to reduce the number of phenomenological coefficients. For example, in crystals it can be shown that $\overleftrightarrow{\lambda}$, the heat conduction tensor, is symmetric in nineteen out of the thirty-two possible point symmetries and in the others, it is of a simple form, by using the crystalline symmetries only (cf. Miller [7]).

As a consequence of the fact that different tensorial phenomena do not couple, it is found that the entropy production separates into several terms (for isotropic media), each of which is separately positive semi-definite.

For example, if

$$\sigma_s = \sum_i X_i Y_i + \sum_j \vec{X}_j \cdot \vec{Y}_j + \sum_k \overleftrightarrow{X}_k : \overleftrightarrow{Y}_k \geq 0 \qquad (3.8)$$

where X_i, Y_i are scalars, $\vec{X}_j$, $\vec{Y}_j$ are vectors, and $\overleftrightarrow{X}_k$, $\overleftrightarrow{Y}_k$ are tensors, then, if the fluxes are linear functions of the forces, it may be concluded that

$$\begin{aligned} &\sum_i X_i Y_i \geq 0 \\ &\sum_j \vec{X}_j \cdot \vec{Y}_j \geq 0 \\ &\sum_k \overleftrightarrow{X}_k : \overleftrightarrow{Y}_k \geq 0 \quad . \end{aligned} \qquad (3.9)$$

In addition to these restrictions which are due to the spatial symmetries of the system, there is another class of symmetry relations between the coefficients L_{ik}, referred to as Onsager reciprocal relations.

§3.2 The Onsager Reciprocal Relations

The Onsager reciprocal relations may be stated as

$$L_{ik} = L_{ki} \quad . \qquad (3.10)$$

This means, for example, that the heat conductivity tensor is symmetric for all materials, not just for special crystalline symmetries. For other explicit examples, see Miller [7].

A derivation of this theorem, along with a discussion of the assumptions implicit in the derivation, will follow. Also, a short discussion of the experimental evidence for the Onsager relations will be given. For now, it will merely be noted that the Onsager relations are generally considered to result from microscopic reversibility of the equations of motion. Some modification of the form of the Onsager relations is necessary in the case where the equations of motion are not simply time reversal invariant This occurs, for example, if magnetic fields or coriolis forces are present, since these change sign under time reversal.

Before giving a derivation of the Onsager relations, some discussion of the effects of different choices of forces and fluxes upon the Onsager relations will be presented.

In the derivation of the Onsager relations, it is assumed that both the forces X_i and the fluxes Y_i separately form linearly independent sets. In choosing appropriate forces and fluxes, linear dependence might occur in two ways. First, intuition might be so faulty that what was believed to be an independent force or flux was merely a linear combination of several other forces or fluxes already considered. More likely, however, is that some constraint upon the system would be neglected. For example, if, in multicomponent diffusion, the various mass currents $\vec{j}_\gamma$ are defined with respect to the center of mass, there would be a constraint that $\sum_{\gamma=1}^{N} \vec{j}_\gamma = 0$, which would result in dependent fluxes. Similarly, the forces could be made dependent by a constraint that the system be in mechanical equilibrium (cf. Hooyman and de Groot [10]). If, for example, a constraint $\sum_{i=1}^{n} \alpha_i Y_i = 0$ relates the fluxes, while the driving forces, X_i, are independent, then the entropy production may be written as

$$\sigma_s = \sum_{i=1}^{n} Y_i X_i = \sum_{i=1}^{n-1} Y_i \{X_i - \frac{\alpha_i}{\alpha_n} X_n\} \quad . \qquad (3.11)$$

With the new set of (n-1) independent forces and fluxes and with phenomenological equations of the form

$$Y_j = \sum_{i=1}^{n-1} \ell_{ji} (X_i - \frac{\alpha_i}{\alpha_n} X_n) \quad , \qquad (3.12)$$

if there are Onsager relations of the form

$$\ell_{ij} = \ell_{ji} \qquad (3.13)$$

and if a matrix L_{ij} is defined by

$$Y_j = \sum_{i=1}^{n} L_{ji} X_i \quad , \qquad (3.14)$$

then
$$L_{ij} = L_{ji} \quad .$$

That is, even for the non-independent set Y_i (i=1,2,...,n), the Onsager relations hold. Of course, there are additional constraints on the coefficients L_{ij} of the form

$$\sum_{i=1}^{n} \alpha_j L_{ij} = 0 \quad . \qquad (3.15)$$

Similarly, if the forces are related by a constraint while the fluxes are independent, the Onsager relations remain valid.

However, in the general case where both the forces and the fluxes are dependent, no Onsager relation holds. Under such conditions, there is a certain arbitrariness in the choice of phenomenological coefficients in the dependent sets. The coefficients L_{ij} can be chosen such that $L_{ij} = L_{ji}$, but need not be. Hence, the validity of the Onsager theorem is destroyed and, to ensure the Onsager theorem's validity, a complete and independent set of fluxes or forces must be chosen.

It can also be shown that, once the fluxes and forces are chosen and Onsager relations established for them, these relations remain valid even if a linear transformation of the bases is made. These transformations are restricted by the condition that, if the forces are transformed, the value of σ_s must be preserved. This invariance of σ_s is also the condition which ensures that L_{ij} and L_{ji} have the same units. That is, if [Y] denotes the dimension of Y, then if $[Y_1] = \beta_1$, $[Y_2] = \beta_2$, $[X_1] = \alpha_1$, and $[X_2] = \alpha_2$,

$$[Y_1] = [L_{11}][X_1] + [L_{12}][X_2]$$

$$[Y_2] = [L_{21}][X_1] + [L_{22}][X_2] \quad . \qquad (3.16)$$

Therefore,

$$[L_{12}] = \beta_1/\alpha_2$$

and
$$[L_{21}] = \beta_2/\alpha_1 \quad . \tag{3.17}$$

However,
$$[Y_1][X_1] = [Y_2][X_2] = \frac{\text{entropy}}{\text{volume-time}} \tag{3.18}$$

so
$$\beta_1 \alpha_1 = \beta_2 \alpha_2$$

or
$$\beta_1/\alpha_2 = \beta_2/\alpha_1 \quad . \tag{3.19}$$

Thus L_{12} and L_{21} have the same dimensions.

To see the transformation properties of Onsager relations, consider the non-singular linear transformation of the forces

$$X_i' = \sum_j A_{ij} X_j \quad . \tag{3.20}$$

Assume also that the fluxes obey linear phenomenological equations

$$Y_i = \sum_j L_{ij} X_j \tag{3.21}$$

where $L_{ij} = L_{ji}$.

In order for σ_s to be invariant, the fluxes must also be transformed to

$$Y_i' = \sum_j B_{ij} Y_j \tag{3.22}$$

where B_{ij} is to be determined from

$$\sigma_s = \sum_i X_i \; Y_i = \sum_i X'_i \; Y'_i \qquad (3.23)$$

or $$\sum_{i,j} L_{ij} \; X_i \; X_j = \sum_{i,j,k,m} A_{ij} \; B_{ik} \; L_{km} \; X_j \; X_m \quad . \qquad (3.24)$$

The most general solution of eqn. (3.24) is

$$\sum_{i,k} A_{ij} \; B_{ik} \; L_{km} = L_{jm} + D_{jm} \qquad (3.25)$$

where D_{jm} is an arbitrary antisymmetric matrix ($D_{jm} = - D_{mj}$). Then the phenomenological equations become

$$Y'_i = \sum_j L'_{ij} \; X'_j \qquad (3.26)$$

where

$$L'_{ij} = \sum_{k,\ell} B_{ik} \; L_{k\ell} \; A^{-1}_{\ell j} = \sum_{k,\ell} (A^{-1}_{ki} \; L_{k\ell} \; A^{-1}_{\ell j}) + \sum_{k,\ell} (A^{-1}_{ki} \; D_{k\ell} \; A^{-1}_{\ell j}) \quad . \qquad (3.27)$$

Also,

$$A_{ij} = B^{-1}_{ji} + \sum_{m,n} D_{im} \; L^{-1}_{mn} \; B^{-1}_{nj} \quad . \qquad (3.28)$$

If the transformations of the forces are now restricted to those which are independent of the coefficients L_{ij}, then $D_{k\ell} = 0$ and

$$L'_{ij} = \sum_{k,\ell} A^{-1}_{\;i} \; L_{k\ell} \; A^{-1}_{\ell j} \quad . \qquad (3.29)$$

Therefore, $L_{ij} = L_{ji}$ implies

$$L'_{ij} = L'_{ji} \quad . \tag{3.30}$$

Davies [11] argues that transformations involving L_{ij} are unphysical. However, he gives no explanation of his claim and later de Groot, Hooyman, and Mazur [12] and also Truesdell [9] argue that exclusion of such transformations cannot be justified on this basis. A transformation of the forces which involves L_{ij} implicitly involves the fluxes. Truesdell [9] gives an example where the original sets are $(X_1, X_2, \ldots, X_n)$, $(Y_1, Y_2, \ldots, Y_n)$ and the transformed sets are $(Y_1, X_2, \ldots, X_n)$ and $(X_1, Y_2, \ldots, Y_n)$. Under such transformations (equivalent to transformations of the forces involving L_{ij}), the Onsager relations do not remain valid in general. This makes it even more desirable to have an explicit rule for separating forces and fluxes.

In Casimir's [13] derivation of the Onsager relations, it is assumed that the deviation of entropy from its equilibrium value is a negative semi-definite quadratic functional of the deviations from their equilibrium values of the thermodynamic variables. That is, if the thermodynamic variables are x^i $(\equiv x_0^i + \alpha^i)$ and if $\Delta S = S - S_0$ where x_0^i and S_0 are the equilibrium values of x^i and S, then

$$\Delta S = -\frac{1}{2} \sum_{i,k} S_{ik}\, \alpha^i \alpha^k \tag{3.31}$$

where S_{ik} is a positive semi-definite form.

Then, it is assumed that the fluxes are the time derivatives of the α^i's. This would seem to give a criterion for choosing the forces and fluxes, but Truesdell [9] argues firstly that merely being the time derivative of "something" is meaningless so a prescription for choosing the appropriate variables α^i is still needed. He also raises a second objection, namely that in the cases where Onsager relations are usually applied, it is not at all obvious that the fluxes under consideration are the time derivatives of any significant quantity. In Casimir's first example of diffusion of gas through a hole in a partition separating gases at two different temperatures, the numbers n_1 and n_2 of molecules in each compartment are obvious choices of variables and, using $\frac{\partial n_1}{\partial t}$, $\frac{\partial n_2}{\partial t}$, internal energy, and its time derivative, it is possible to correctly derive Onsager relations between the concentration difference caused by the differing temperatures and the energy carried by the molecules. However, in the more general cases to which Onsager relations are often applied, for instance in deriving the symmetry of the heat conduction tensor, it is necessary to use more complicated and less convincing arguments to show that the derivation of the Onsager relations is still applicable (cf. Truesdell [9] or Casimir [13]). Truesdell argues that in other thermodynamic processes, there is even less justification for assuming the Onsager relations to be valid.

It has been shown that the invariance of σ_s alone is insufficient to guarantee the validity of the Onsager relations under general linear transformations of forces and fluxes. Hooyman, de Groot and Mazur [12] propose that a more restrictive condition be applied. If it is assumed that the entropy with respect to equilibrium, rather than the entropy production σ_s is required to be invariant under the transformations, then the symmetry of L_{ij} is preserved. However, this criterion suffers from the drawback that it has never been applied to show the validity of the Onsager relations in common thermodynamic situations. This is probably due to the fact that ΔS is expressed as a quadratic functional of the α^i's rather than directly in terms of the usual forces and fluxes, making application difficult.

As far as the experimental evidence is concerned, there exists a comprehensive review article by Miller [7]. His uncritical conclusion is that, for all the phenomena he considers, what evidence there is agrees with the Onsager reciprocal relations. Truesdell's [9] more critical analysis is that there is insufficient data in several cases, notably with respect to symmetry of the heat conduction tensor, and that, in other cases, particularly for isothermal ternary diffusion, the limited data seems to contradict Onsager's theorem. However, due to the difficulty of the experiments and to the fact that they have not been repeated, it is possible that the apparent disagreements would be removed by more accurate experiments. It would seem that there is no definite counterexample to the Onsager relations and that, for many phenomena, there is good experimental evidence for their validity. However, in many cases, the symmetry of L_{ij} can be derived independently from Onsager's theorem and hence these cases do not serve to confirm the Onsager theorem in all its generality.

In order to derive the Onsager relations, it is useful to first consider a transparent case, that of a simple chemical reaction triangle (due to Onsager [14]). That is, there are three chemical reactants A, B, and C with n_A, n_B, and n_C molecules respectively; if $\overline{n}_A$, $\overline{n}_B$, and $\overline{n}_C$ are their equilibrium numbers, and if the rate at which species A spontaneously transforms into species B is κ_{BA} (a constant), the equations for the numbers of molecules of the various species are

$$\frac{dn_A}{dt} = -(\kappa_{BA} + \kappa_{CA})n_A + \kappa_{AB}n_B + \kappa_{AC}n_C$$

$$\frac{dn_B}{dt} = \kappa_{BA}n_A - (\kappa_{AB} + \kappa_{CB})n_B + \kappa_{BC}n_C$$

$$\frac{dn_C}{dt} = \kappa_{CA}n_A + \kappa_{CB}n_B - (\kappa_{AC} + \kappa_{BC})n_C \qquad (3.32)$$

subject to the conditions

$$n_A + n_B + n_C = \bar{n}_A + \bar{n}_B + \bar{n}_C = n \qquad (3.33)$$

and

$$\left.\frac{dn_A}{dt}\right|_{\bar{n}_A} = \left.\frac{dn_B}{dt}\right|_{\bar{n}_B} = \left.\frac{dn_C}{dt}\right|_{\bar{n}_C} = 0 \quad . \qquad (3.34)$$

If the equilibrium concentrations are known, then there are two independent constraints on the six coefficients κ_{AB}, κ_{AC}, ..., κ_{CB} . Thus, even if the system is assumed to be in equilibrium, there are still four degrees of freedom in choosing the rate coefficients.

However, the usual procedure in chemistry is to assume detailed balance. That is, it is assumed that, in equilibrium, each simple reaction exactly balances itself (i.e., $\kappa_{BA}\bar{n}_A = \kappa_{AB}\bar{n}_B$, etc.). It is possible to imagine equilibrium situations not obeying detailed balance (e.g., $\kappa_{AC} = \kappa_{BA} = \kappa_{CB} = 2$; $\kappa_{CA} = \kappa_{AB} = \kappa_{CB} = 1$). Such a reaction, while it satisfies the equilibrium requirement, is ruled out by considering time reversal. In the simplest physical situations (zero magnetic field, etc.), the underlying microscopic laws of motion are time reversal invariant. Thus, for any microstate there is a corresponding state with all velocities reversed. Since in equilibrium all allowable states have equal a priori probabilities, each time reversed motion occurs, on the average, as often as the original motion. As a consequence, in equilibrium, each direct reaction A → B must occur as often as the reverse reaction B → A. This means that detailed balance holds as a consequence of the time reversal invariance of the microscopic motions. The consequences of such time reversal invariance were pointed out much earlier by Nernst (1883) and Bohr (1913).

The analogy between linear rate equations in chemical reactions and the general phenomenological equations discussed earlier can be made closer by rewriting the rate equations in terms of forces and fluxes.

For an ideal solution at constant pressure and temperature, the free energy G can be written as

$$G = G_{equil} + RT \{n_A \ln (\frac{n_A}{\bar{n}_A}) + n_B \ln (\frac{n_B}{\bar{n}_B}) + n_C \ln (\frac{n_C}{\bar{n}_C})\} \tag{3.35}$$

where G_{equil} is the equilibrium (minimum) value of G. Then, to first order in δn_A, δn_B, δn_C (small changes in n_A, n_B, n_C)

$$\delta G \Big|_{p,T,n} = RT \{\delta n_A \ln(\frac{n_A}{\bar{n}_A}) + \delta n_B \ln(\frac{n_B}{\bar{n}_B}) + \delta n_C \ln(\frac{n_C}{\bar{n}_C})\} \; . \tag{3.36}$$

Defining $\alpha_A = n_A - \bar{n}_A$, etc. gives the fluxes as $\dot{\alpha}_A = \dot{n}_A$ (the dot denotes time derivative), and also

$$\delta G \Big|_{p,T,n} = - X_A \, \delta\alpha_A - X_B \, \delta\alpha_B - X_C \, \delta\alpha_C \quad . \tag{3.37}$$

The "forces" X_A, X_B, and X_C are now seen to be

$$X_A = - RT \ln(\frac{n_A}{\bar{n}_A}) \approx - \frac{RT}{\bar{n}_A} \alpha_A, \text{ etc.} \tag{3.38}$$

for small α_A, α_B, α_C .

The rate equations can be written in terms of the forces X_A, X_B, and X_C, and the fluxes $\dot{\alpha}_A$, $\dot{\alpha}_B$, and $\dot{\alpha}_C$ as

$$\begin{aligned}
\dot{\alpha}_A &= (\kappa_{BA} + \kappa_{CA}) \frac{\bar{n}_A}{RT} X_A - \frac{\kappa_{AB} \bar{n}_B}{RT} X_B - \frac{\kappa_{AC} \bar{n}_C}{RT} X_C \\
\dot{\alpha}_B &= - \frac{\kappa_{BA} \bar{n}_A}{RT} X_A + (\kappa_{AB} + \kappa_{CB}) \frac{\bar{n}_B}{RT} X_B - \frac{\kappa_{BC} \bar{n}_C}{RT} X_C \\
\dot{\alpha}_C &= - \frac{\kappa_{CA} \bar{n}_A}{RT} X_A - \frac{\kappa_{CB} \bar{n}_B}{RT} X_B + (\kappa_{AC} + \kappa_{BC}) \frac{\bar{n}_C}{RT} X_C \quad .
\end{aligned} \tag{3.39}$$

This is the form suitable for the application of Onsager's theorem and it is seen to be valid, due to micro-reversibility.

The next step is to generalize the above considerations in order to derive Onsager relations for any thermodynamic forces and fluxes in the near equilibrium (linear) range. In the example of a chemical triangle reaction, the reactions A → B, B → C, etc. were considered. These reactions might be thought of as progressing with non-zero rates κ_{BA}, κ_{CB}, etc., even in equilibrium. Although in equilibrium the rates balance in such a way that the net rates are zero, at a microscopic level the separate forward and reverse reaction rates retain a clear intuitive meaning. It was considerations of micro-reversibility which made it possible to derive Onsager relations for the observed rates near equilibrium. In order to generalize this derivation, an analogous quantity to the rates is needed. This quantity must be time reversal invariant on the average. The most obvious choice which is common to all thermodynamic systems is to consider fluctuations of thermodynamic variables near to equilibrium.

In deriving Onsager's theorem, the method due to Casimir [13] will be followed. Due to the assumption of local equilibrium, the entropy S may be expanded as a Taylor series about its equilibrium value. Writing $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ to denote the deviations of the extensive thermodynamic variables A_i from their equilibrium values A_i^{eq} (i.e., $\alpha_i = A_i - A_i^{eq}$), gives $\vec{\alpha} = 0$ in equilibrium and

$$S(\vec{\alpha}) = S(o) + \sum_{i=1}^{n} \frac{\partial S}{\partial \alpha_i}\bigg|_{\vec{\alpha}=0} \alpha_i + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 S}{\partial \alpha_i \partial \alpha_j}\bigg|_{\vec{\alpha}=0} \alpha_i \alpha_j + \ldots \quad . \tag{3.40}$$

Since S is a maximum in equilibrium,

$$\frac{\partial S}{\partial \alpha_i}\bigg|_{\vec{\alpha}=0} = 0 \tag{3.41}$$

and therefore

$$\Delta S(\equiv S(\vec{\alpha}) - S(o)) = -\frac{1}{2}\sum_{i,j=1}^{n} S_{ij}\,\alpha_i\alpha_j \leq 0 \quad (3.42)$$

where

$$S_{ij} = S_{ji} = -\left.\frac{\partial^2 S}{\partial\alpha_i\partial\alpha_j}\right|_{\vec{\alpha}=0} \quad . \quad (3.43)$$

Since entropy is an extensive variable, as are the α_i's, $\sum_{j=1}^{n} S_{ij}\,\alpha_j$ is an intensive variable.

The conjugate driving forces X_i may be defined, in the spirit of equilibrium thermodynamics, as

$$X_i = \left(\frac{\partial S}{\partial\alpha_i}\right) \approx \sum_{j=1}^{n} S_{ij}\,\alpha_j \quad . \quad (3.44)$$

As local equilibrium has been assumed, it is also assumed that expectation values of the near equilibrium variables may be defined as averages over an ensemble. For this, the Einstein formula for the probability density $W(\vec{\alpha})$ is used:

$$W(\vec{\alpha})d\vec{\alpha} = \frac{e^{k\Delta S}\,d\vec{\alpha}}{\int e^{k\Delta S}\,d\vec{\alpha}} \quad . \quad (3.45)$$

With this expression for the probability density, several averages can be computed (cf. de Groot and Mazur [4] pp. 87-89). Assuming the α_i are independent,

$$\langle\alpha_i\ X_j\rangle_{AV} = -\,k\,\delta_{ij} \quad .$$

Then, since

$$\alpha_i = \sum_{j=1}^{n} S_{ij}^{-1} X_j \quad , \tag{3.46}$$

$$\langle \alpha_i \alpha_j \rangle_{AV} = k S_{ij}^{-1} \quad . \tag{3.47}$$

For simplicity, only variables which are even under time reversal will be considered initially (e.g., kinetic energies). Such variables as angular velocities and such phenomena as magnetic fields which change sign under time reversal will be excluded for now.

Since for every microscopic motion there is an equally probable motion with all velocities reversed, time reversal invariance holds in the sense that the average of a variable α_i must be the same for equal positive and negative times from some initial instant. That is,

$$\langle \alpha_i(t + \tau) \rangle_{AV}^{\vec{\alpha}(t)} = \langle \alpha_i(t - \tau) \rangle_{AV}^{\vec{\alpha}(t)} \tag{3.48}$$

where the $\vec{\alpha}(t)$ superscript denotes the fact that the average is now over all states which have the initial condition $\vec{\alpha} = \vec{\alpha}(t)$ at t . If this is now multiplied by $\alpha_j(t)$ and averaged over all possible initial conditions $\vec{\alpha}(t)$, the result being denoted by $\langle \alpha_j(t)\ \alpha_i(t + \tau) \rangle$, then

$$\langle \alpha_j(t)\ \alpha_i(t + \tau) \rangle = \langle \alpha_j(t)\ \alpha_i(t - \tau) \rangle \quad . \tag{3.49}$$

This is a statement of the assumption of microscopic reversibility, valid near equilibrium.

If the system were exactly in equilibrium, the result would be independent of t and it may be assumed that

$$<\alpha_j(t)\ \alpha_i(t - \tau)> = <\alpha_j(t + \tau)\ \alpha_i(t)> \qquad (3.50)$$

provided that $\tau << \tau^{ev}$ where τ^{ev} is a macroscopic evolution time for the system as a whole.

If the thermodynamic fluxes in the phenomenological equations are considered to be $\dot{\alpha}_i$, the time derivatives of some thermodynamic variables, then the consequences of time reversal invariance upon the phenomenological coefficients may be seen. In order to derive Onsager relations, some assumption must be made about regression of fluctuations (i.e., about the form of the equations of motion for α_i). The macroscopic linear equations are of the form $\dot{\alpha}_i = \sum_j L_{ij} X_j$. The usual procedure (cf. Onsager [14]) is to apply this same macroscopic equation to the microscopic fluctuations α_i. Although the linearity of this equation may make this procedure seem plausible, it is possible to imagine a psuedo-linearity which holds only at moderately large values of $\vec{\alpha}$. Also, as will be seen, there is some difficulty in defining the time derivative of α_i .

In order to extend the linear laws to the range of fluctuations, a Langevin equation could be substituted for the simple linear equation. That is, consider the equation

$$\dot{\alpha}_i = \sum_j L_{ij}\ X_j + \kappa_i(t) = \sum_j P_{ij}\ \alpha_j + \kappa_i(t) \qquad (3.51)$$

where $P_{ij} = \sum_{\ell,k} L_{ik} S_{kj}$ and $\langle \kappa_i(t) \rangle = 0$ for time averaging over times much greater than τ_1, the microscopic relaxation time for the system. This is simply adding a fluctuating force $\vec{\kappa}$ to the linear force. This equation may be solved for $\alpha(t)$ as

$$\vec{\alpha}(t) = e^{t\overleftrightarrow{P}} \vec{\alpha}(o) + e^{t\overleftrightarrow{P}} \int_o^t e^{-t'\overleftrightarrow{P}} \vec{\kappa}(t')dt' \quad . \quad (3.52)$$

Thus,

$$\langle \vec{\alpha}(t + \tau) - \vec{\alpha}(t) \rangle^{\alpha(t)}_{AV} = e^{t\overleftrightarrow{P}} [e^{\tau\overleftrightarrow{P}} - 1] \vec{\alpha}(o) \quad . \quad (3.53)$$

The macroscopic linear equations would only be expected to hold at a microscopic level for $\tau >> \tau_1$. Then, if $\tau\overleftrightarrow{P} << 1$ (i.e., $\tau << \tau^{ev}$), it is valid to write

$$\langle \vec{\alpha}(t + \tau) - \vec{\alpha}(t) \rangle^{\vec{\alpha}(t)}_{AV} \approx \tau P e^{t\overleftrightarrow{P}} \vec{\alpha}(o) = \tau\overleftrightarrow{P} \langle \vec{\alpha}(t) \rangle$$

$$= \tau \sum_j L_{ij} \langle X_j \rangle \quad . \quad (3.54)$$

Therefore,

$$\langle \alpha_\ell(t)[\alpha_k(t + \tau) - \alpha_k(t)] \rangle = \tau \sum_j L_{kj} \langle \alpha_\ell X_j \rangle$$

$$= k \tau L_{k\ell} \quad . \quad (3.55)$$

Similarly,

$$\langle \alpha_k (t)[\alpha_\ell(t + \tau) - \alpha_\ell(t)] \rangle = k \tau L_{\ell k} \quad .$$

However, microscopic reversibility implies that

$$\begin{aligned}&\langle\alpha_\ell(t)[\alpha_k(t+\tau)-\alpha_k(t)]\rangle\\ &\quad=\langle\alpha_\ell(t)[\alpha_k(t-\tau)-\alpha_k(t)]\rangle\\ &\quad=\langle\alpha_\ell(t+\tau)[\alpha_k(t)-\alpha_k(t+\tau)]\rangle\\ &\quad=\langle\alpha_k(t)[\alpha_\ell(t+\tau)-\alpha_\ell(t)]\rangle\quad.\end{aligned}\tag{3.56}$$

This completes the derivation of the Onsager reciprocal relations

$$L_{\ell k}=L_{k\ell}\quad.\tag{3.57}$$

However, there are several steps in the derivation which qualify only as hypotheses. Some of these are:

(1) near equilibrium, ΔS is a quadratic function of the α_i's;

(2) entropy is proportional to the log of the volume in Γ-space near equilibrium (These first two are essentially the assumption of local equilibrium.);

(3) in the fluctuation range, the forces and fluxes obey a linear law (or a Langevin equation) and, more importantly, that the coefficients in this linear law are identical to the coefficients in the macroscopic phenomenological law;

(4) the system has a relaxation time τ_1 in which it reaches a steady state, such that $\tau_1 << \tau^{ev}$ where τ^{ev} is the macroscopic evolution time of the system; and

(5) the forces and fluxes considered in a real system are actually of the form assumed in the derivation, or can be reduced to such a form.

The assumption that the variables α_i are even under time reversal is not essential and the result can be generalized. If β_i are variables which change sign under time reversal (e.g., angular momentum) then

$$<\beta_\lambda(t)\beta_\mu(t + \tau)> = <\beta_\lambda(t)\beta_\mu(t - \tau)>$$

and

$$<\alpha_i(t)\beta_\mu(t + \tau)> = - <\alpha_i(t)\beta_\mu(t - \tau)> \quad . \tag{3.58}$$

Entropy, being an even function, can be written

$$\Delta S = - \frac{1}{2} \{ \sum_{i,j=1}^{n} S_{ij}\, \underbrace{\alpha_i\alpha_j}_{} + \sum_{\lambda,\mu=n+1}^{m} S_{\lambda\mu}\beta_\lambda\beta_\mu \} \tag{3.59}$$

and the phenomenological equations can be written as

$$\dot{\alpha}_i = \sum_{j=1}^{n} L_{ij}X_j + \sum_{\lambda=n+1}^{m} L_{i\lambda}X_\lambda \quad (i=1,2,\ldots n)$$

$$\dot{\beta}_\lambda = \sum_{j=1}^{n} L_{\lambda j}X_j + \sum_{\mu=n+1}^{m} L_{\lambda\mu}X_\mu \quad (\lambda=n+1,\ldots m) \tag{3.60}$$

where

$$X_i = \sum_{j=1}^{n} S_{ij}\alpha_j$$

$$X_\lambda = \sum_{\mu=n+1}^{m} S_{\lambda\mu}\beta_\mu \quad . \qquad (3.61)$$

The corresponding generalized Onsager relations are

$$L_{ij} = L_{ji}$$

$$L_{\lambda\mu} = L_{\mu\lambda}$$

$$L_{i\lambda} = - L_{\lambda i} \quad . \qquad (3.62)$$

Also, if magnetic fields are present, the motions are the same if all the particle motions and the magnetic fields are simultaneously reversed. Hence,

$$L_{ij}(\vec{B}) = L_{ji}(-\vec{B})$$

$$L_{\lambda\mu}(\vec{B}) = L_{\mu\lambda}(-\vec{B})$$

$$L_{i\lambda}(\vec{B}) = - L_{\lambda i}(-\vec{B}) \qquad (3.63)$$

is the general form of the Onsager relations where $\vec{B}$ represents magnetic fields (or any other forces odd under time reversal).

Thus, it is seen that, while the Onsager relations do have a basis in microscopic theory, they are far from being rigorously proven in all their generality. It seems the assumptions necessary to derive the Onsager relations place restrictions on the type of system to which they are applicable. Casimir [13] argues, and it appears reasonable, that the Onsager relations should apply to systems which can be treated by the Boltzmann equation.

For clarity, a straight-forward example of Onsager relations is presented. Consider a container separated into two equal compartments by a wall with a small hole (due to Casimir [13]). Let n_1 and n_2 be the number of moles of an ideal gas in each compartment, and T_1 and T_2 be their respective temperatures.

In equilibrium

$$\begin{aligned} n_1 &= n_2 = n \\ T_1 &= T_2 = T \quad . \end{aligned} \tag{3.64}$$

For an ideal gas, the energy u is proportional to $n_1 T_1$ so conservation of particle number and of energy gives the following conditions on the fluctuations δn_1, δn_2, δT_1, and δT_2:

$$\begin{aligned} \delta n_1 + \delta n_2 &= 0 \\ \delta(n_1 T_1 + n_2 T_2) &= 0 \quad . \end{aligned} \tag{3.65}$$

Eqns. (3.65) imply

$$(\delta T_1 + \delta T_2) = \frac{\delta n_2}{n} (\delta T_1 - \delta T_2) \tag{3.66}$$

and, to this order of accuracy

$$(\delta T_1)^2 = (\delta T_2)^2 \quad . \tag{3.67}$$

For an ideal gas,

$$S = nC_v \ln T + R \ln V - nR \ln n \quad . \qquad (3.68)$$

Manipulating yields

$$\Delta S = - \frac{nC_v}{T^2} (\delta T_2)^2 - \frac{R}{n} (\delta n_2)^2 \qquad (3.69)$$

to first non-vanishing order in the fluctuations.

Transforming to new variables α_1 and α_2 and conjugate forces X_1 and X_2 where

$$\alpha_1 = \delta n_2$$

$$\alpha_2 = \delta u_2 = (T\delta n_2 + n\delta T_2)C_v \qquad (3.70)$$

gives

$$\Delta S = - \frac{(C_v + R)}{n} \alpha_1^2 - \frac{1}{nT^2 C_v} \alpha_2^2 + \frac{2C_v}{nT} \alpha_1\alpha_2 \qquad (3.71)$$

and $$X_1 = \frac{\partial(\Delta S)}{\partial\alpha_1} = 2 \left(\frac{R}{n} \delta n_2 - \frac{C_v}{T} \delta T_2\right)$$

$$X_2 = \frac{\partial(\Delta S)}{\partial\alpha_2} = \frac{2\delta T_2}{T^2} \quad . \qquad (3.72)$$

If δn_2 and $\frac{\delta T_2}{T^2}$ are regarded as the forces causing the mass and energy flows $\delta\dot{n}_2$ and $\delta\dot{u}_2$, then the linear phenomenological equations are

$$\delta\dot{n}_2 = A\ \delta u_2 + B \frac{\delta T_2}{T^2}$$

$$\delta\dot{u}_2 = Q\ \delta n_2 + \omega \frac{\delta T_2}{T^2} \quad . \qquad (3.73)$$

In eqns. (3.73), Q is a heat convection coefficient and ω is a heat conduction coefficient. If, in order to apply Onsager relations, eqns. (3.73) are put in terms of α_1, α_2, X_1, and X_2 then

$$\dot{\alpha}_1 = (\frac{1}{2}\frac{An}{R})X_1 + \frac{1}{2}(\frac{AnTC_v}{R} + B)X_2$$
$$\dot{\alpha}_2 = (\frac{1}{2}\frac{AnQ}{R})X_1 + \frac{1}{2}(\frac{AnTC_vQ}{R} + BQ + \omega)X_2 \qquad (3.74)$$

and the Onsager relation is

$$\frac{B}{A} = \frac{(Q - TC_v)}{R}\, n \quad . \qquad (3.75)$$

If a steady state is established ($\delta\dot{n} = 0$) while a temperature difference is maintained, then

$$\delta n = -\frac{B}{A}\frac{\delta T}{T^2} \qquad (3.76)$$

or

$$\frac{\delta n}{n} = -(Q/T - C_v)\frac{\delta T}{RT} \quad . \qquad (3.77)$$

Eqn. (3.77) relates the concentration difference caused by unequal temperatures, to the energy carried by the molecules. The value of Q may be calculated from kinetic theory. If ϕ is the diameter of the hole and λ is the mean free path of the molecules, then there are two limiting cases.

(1) If $\lambda >> \phi$, then $Q = (C_v + \frac{1}{2}R)T$

so

$$\frac{\delta n}{n} = -\frac{1}{2}\frac{\delta T}{T}$$

or

$$n = -2T\frac{\delta n}{\delta T} \quad . \qquad (3.78)$$

Eqn. (3.78) results in Knudsen's formula

$$n(T) = \frac{\text{constant}}{T^{1/2}} \quad . \tag{3.79}$$

(2) If $\lambda << \phi$, then $Q = (C_v + R)T$

so

$$\frac{\delta n}{n} = - \frac{\delta T}{T}$$

or

$$n = - T \frac{\delta n}{\delta T} \tag{3.80}$$

which yields

$$n(T) = \frac{\text{constant}}{T} \quad . \tag{3.81}$$

In his 1945 paper, Casimir [13] also discusses heat conduction in an anisotropic crystal. In this and many other common systems, the relevant forces and fluxes are not such simple quantities as $\delta\dot{n}$ and $\delta\dot{u}$, but rather field variables (e.g., $\vec{\nabla}T(\vec{r},t)$). This necessitates the consideration of time derivatives of integrals of field variables over regions of space. In the example of heat conduction, this results in an Onsager relation weaker than the statement that the heat conduction tensor is symmetric. Instead, it may be concluded that, if $\overleftrightarrow{\lambda} = \overleftrightarrow{\lambda}^{(s)} + \overleftrightarrow{\lambda}^{(a)}$ where $\overleftrightarrow{\lambda}^{(a)}$ is the antisymmetric part of $\overleftrightarrow{\lambda}$, then

$$\vec{\nabla} \cdot \overleftrightarrow{\lambda}^{(a)} = 0 \quad . \tag{3.82}$$

Truesdell [9] points out that in a homogeneous system this implies

$$\frac{\partial \overleftrightarrow{\lambda}^{(a)}}{\partial T} \cdot (\vec{\nabla} T) = 0 \tag{3.83}$$

for any $T(\vec{r},t)$ whatever. Thus, $\overleftrightarrow{\lambda}^{(a)}$ is constant. Casimir [13] argues that it is only $\vec{\nabla} \cdot \overleftrightarrow{\lambda}$ that is of physical importance and so that it is permissible to set $\overleftrightarrow{\lambda}^{(a)} = 0$, which gives the Onsager relation

$$\overleftrightarrow{\lambda}_{\alpha\beta} = \overleftrightarrow{\lambda}_{\beta\alpha} \quad . \tag{3.84}$$

While it is true that the differential equation for $T(\vec{r},t)$ involves only $\vec{\nabla} \cdot \overleftrightarrow{\lambda}$, it is still possible to have boundary conditions which involve $\overleftrightarrow{\lambda}^{(a)}$ itself. For example, at a thermally insulated boundary, an appropriate boundary condition is

$$\hat{n} \cdot [(\overleftrightarrow{\lambda}^{(a)} + \overleftrightarrow{\lambda}^{(s)}) \cdot (\vec{\nabla} T)] = 0 \tag{3.85}$$

where $\hat{n}$ is a unit vector normal to the surface. An experiment which would in fact determine $\overleftrightarrow{\lambda}^{(a)}$ is described by Truesdell [9], p. 114. In fact, the classic heat experiments of Soret and Voight claim to measure $\overleftrightarrow{\lambda}$ itself.

Onsager [13], however, claims to show that $\overleftrightarrow{\lambda}^{(a)} = 0$ directly, in agreement with the experimental evidence. However, as he considers the entire crystal rather than a microscopic volume of it, it is not obvious how the Onsager relation he derives is valid for the local heat conduction tensor in its usual interpretation. The situation and the parameters Onsager considers do not generalize easily to the experimental situations of Soret, Voight, etc.

§3.3 Minimum Entropy Production

In the theory of irreversible processes, there is an important class of processes which plays a role analogous to that of equilibrium states in reversible thermodynamics. These are the steady state processes which are characterized by time independent forces and fluxes. This situation differs from an equilibrium state by the fact that, due to the application of constraints, some forces and fluxes are non-zero and hence that the entropy production is non-zero. These stationary non-equilibrium states are sometimes characterized by a minimum of the entropy production (consistent with the constraints), analogous to the maximum entropy of equilibrium states.

Minimum entropy production is valid only for certain conditions and, in order to derive it, it is necessary to assume linear phenomenological laws, the validity of Onsager's theorem, constant phenomenological coefficients, the constraints result from time independent boundary conditions and that there is no mass convection. These assumptions restrict the systems considered to those in which there are no large variations of the parameters (T, ρ, etc.) over the entire system. This is necessary to ensure that the phenomenological coefficients are constant throughout the system.

The simplest way in which to present this theorem is through several examples. First, consider transfer of matter and energy between two phases of a material which are kept at different temperatures. Once steady state is reached, there will be no mass transport ($J_m = 0$) but there will still be heat transfer ($J_q \neq 0$). Then,

$$\sigma_s = X_q J_q + X_m J_m \quad . \tag{3.86}$$

Assuming linear phenomenological laws of obeying Onsager's theorem,

$$\begin{aligned} J_q &= L_{11} X_q + L_{12} X_m \\ J_m &= L_{21} X_q + L_{22} X_m \end{aligned} \tag{3.87}$$

where $L_{12} = L_{21}$, the minimal property of σ_s may be derived.

For steady state, $J_m = 0$ and in order to see the connection with entropy production observe that

$$\left.\frac{\partial \sigma_s}{\partial X_m}\right|_{X_q} = (L_{12} + L_{21}) X_q + 2L_{22} X_m = 2J_m \quad . \tag{3.88}$$

Thus, if σ_s is a minimum (subject to the constraint on X_q), then

$$\left.\frac{\partial \sigma_s}{\partial X_m}\right|_{X_q} = 2J_m = 0 \quad . \tag{3.89}$$

This means that, if the system is in a steady state, then the entropy production is a local minimum (not a maximum since σ_s is a positive semi-definite quadratic form).

A second example is furnished by thermal conduction in an isotropic medium. In this case (cf. de Groot and Mazur [4] p. 45),

$$\sigma_s = \vec{J}_q \cdot \vec{\nabla}(\frac{1}{T}) \qquad (3.90)$$

and the phenomenological equation is

$$\vec{J}_q = L_{qq}\, \vec{\nabla}(\frac{1}{T}) \qquad (3.91)$$

so the total entropy production P is

$$P \equiv \int_V \sigma_s \, dV = \int_V L_{qq}\, (\vec{\nabla}(\frac{1}{T}) \cdot \vec{\nabla}(\frac{1}{T}))dV \quad . \qquad (3.92)$$

The solution of eqn. (3.92) is, from the calculus of variations,

$$\nabla^2\, (\frac{1}{T}) = 0 \qquad (3.93)$$

or

$$\vec{\nabla} \cdot \vec{J}_q = 0 \quad . \qquad (3.94)$$

The local energy equation is

$$\rho \frac{\partial u}{\partial t} = \rho c_v \frac{\partial T}{\partial t} = - \vec{\nabla} \cdot \vec{J}_q \qquad (3.95)$$

so, for a minimum of entropy production,

$$\frac{\partial u}{\partial t} = 0 \qquad (3.96)$$

and there is a steady state since heat conduction is the only process considered.

Note that Onsager relations were not needed to derive the minimum entropy production theorem for heat conduction.

It can also be shown that these steady states with minimum entropy production are stable with respect to small local perturbations. To show this for heat conduction, differentiate P with respect to time to get

$$\frac{\partial P}{\partial t} = 2 \int_V L_{qq} \vec{\nabla}(\frac{1}{T}) \cdot \vec{\nabla}(\frac{\partial}{\partial t} (\frac{1}{T}))dV$$

$$= 2 \int_V \vec{J}_q \cdot \vec{\nabla}(\frac{\partial}{\partial t} (\frac{1}{T}))dV$$

$$= 2 \int_\Sigma (\frac{\partial}{\partial t} (\frac{1}{T}))\vec{J}_q \cdot d\vec{\Sigma} - 2 \int_V (\frac{\partial}{\partial t} (\frac{1}{T}))\vec{\nabla} \cdot \vec{J}_q \, dV . \quad (3.97)$$

With T fixed on the boundary Σ ,

$$\frac{\partial P}{\partial t} = - 2 \int_V \frac{\rho c_v}{T^2} (\frac{\partial T}{\partial t})^2 \, dV \leq 0 \qquad (3.98)$$

since ρ, c_v, and T are positive.

Thus, since P decreases with time, the system returns toward the state of minimum entropy production and hence the steady state is stable.

More general linear phenomena may be considered and similar minimal and stability properties of steady states may be derived (cf. de Groot and Mazur [4]).

It is also possible to derive an evolution criterion without the use of the phenomenological laws. This criterion is not restricted to situations where the coefficients are constant or where the Onsager relations are valid (cf. Glansdorff and Prigogine [15]).

Writing

$$P \equiv \int_V \sigma_s dV = \int_V \sum_i X_i Y_i \, dV \quad . \tag{3.99}$$

then
$$\frac{\partial P}{\partial t} = \int_V \sum_i Y_i \frac{\partial X_i}{\partial t} + \int_V \sum_i \frac{\partial Y_i}{\partial t} X_i \, dV \quad . \tag{3.100}$$

This may be written as

$$\frac{\partial P}{\partial t} = \int_V \left(\frac{\partial_x \sigma_s}{\partial t} + \frac{\partial_y \sigma_s}{\partial t}\right) dV \tag{3.101}$$

where $\frac{\partial_x}{\partial t}$ and $\frac{\partial_y}{\partial t}$ denote the portion of the time derivative which results from the time variation of X_i and Y_i respectively. Then,

$$\frac{\partial_x P}{\partial t} = \int_V \frac{\partial_x \sigma_s}{\partial t} dV = \int_V \sum_i Y_i \frac{\partial X_i}{\partial t} dV \leq 0 \quad . \tag{3.102}$$

The validity of (3.102) is again most easily seen by example, as in de Groot and Mazur [4], p. 54.

This "general evolution criterion" may be interpreted as meaning that the part of the change in entropy production due to the changes of the thermodynamic forces will always be negative, although nothing may be said about the total change in entropy production without more restrictive assumptions.

The question of stability of these states and others will be considered in more detail in the next chapter.

CHAPTER IV

STABILITY AND FLUCTUATIONS

§4.1 Liapounoff Functions and Stability

In considering general thermodynamic states, the subclass of stable states is of obvious interest. In many applications, it is the properties of a system after the transient phenomena have died out which are of interest. Examples of such states, which are known from experience to be stable, are systems in equilibrium and many steady state transport systems. These are stable in the intuitive sense that, if a small change is made in one of the parameters, then, left to itself, the system will not evolve to a state much different from the original state.

The idea of stability can be more precisely formulated in terms of Liapounoff functions. If $\vec{x}$ denotes the independent variables characterizing a system (e.g., volume, energy, entropy, etc.) and if $\vec{x}_0$ is the value of $\vec{x}$ at $t = t_0$ then

$$\vec{x} = \vec{\phi}(t,t_0,\vec{x}_0) \quad . \tag{4.1}$$

Defining

$$\vec{y}(t) = \vec{\phi}(t,t_0,\vec{x}_0 + \vec{\delta}) - \vec{\phi}(t,t_0,\vec{x}_0) \tag{4.2}$$

where $\vec{\delta}$ is some deviation from the state $\vec{x}_0$, then, if for all $\varepsilon > 0$ there corresponds a function, $k(\varepsilon) > 0$, such that

$$|\vec{\phi}(t,t_0,\vec{x}_0 + \vec{\delta}) - \vec{\phi}(t,t_0,\vec{x}_0)| < \varepsilon \qquad (4.3)$$

for all t provided that $|\vec{\delta}| < k(\varepsilon)$, the system is said to be stable (in the sense of Liapounoff).

If also,

$$\lim_{t\to\infty} |\vec{\phi}(t,t_0,\vec{x}_0 + \vec{\delta}) - \vec{\phi}(t,t_0,\vec{x}_0)| = 0 , \qquad (4.4)$$

the system is said to be asymptotically (or completely) stable. Asymptotic stability implies that the perturbed and unperturbed states will evolve toward exactly the same state as $t \to \infty$ (i.e., fluctuations die out).

As a result, if $\frac{\partial}{\partial t}(y^2) \leq 0$, then the system is stable and if $\frac{\partial}{\partial t}(y^2) < 0$ for all $\vec{y}$, the system is asymptotically stable. Choosing another positive definite form rather than y^2 would yield different sufficient conditions for stability.

It is important to notice that the above stability theory applies both to infinitesimal and to finite perturbations $\vec{\delta}$, and is not restricted to steady state situations. However, in many cases, the stability problem cannot be completely solved and stability with respect to infinitesimal perturbations is of interest. If it is shown that a system is not stable with respect to infinitesimal perturbations, then it is completely unstable. However, even if a system is stable with respect to infinitesimal perturbations, it may still be unstable with respect to some finite size of perturbations, in which case the system is called metastable. For example, a particle at the bottom of a finite potential well with finite energy barriers is metastable.

§4.2 Stability of Equilibrium States

Before considering the question of stability for general thermodynamic states, the classical (Gibbs-Duhem) theory of stability of equilibrium states will be considered.

In a reversible process, the first law of thermodynamics states that the heat received by a system, dQ, is

$$dQ = dE + p\ dV \tag{4.5}$$

where dE and dV are the changes in the system's internal energy and volume respectively.

For a closed system at uniform temperature and pressure, the second law may be written as

$$dS = d_e S + d_i S \geq \frac{dQ}{T} \tag{4.6}$$

or alternatively as

$$Td_i S = TdS - dE - p\ dV \geq 0 \tag{4.7}$$

where the equality holds only for completely reversible processes.

If ΔE, ΔS, and ΔV are the small but otherwise arbitrary deviations from the equilibrium values of E, S, and V, then the following condition for the stability of the equilibrium state is deduced:

$$\Delta E + p\,\Delta V - T\,\Delta S > 0 \quad . \qquad (4.8)$$

That is, if no state which starts at equilibrium can change in such a way as to satisfy the second law, then the system must remain in equilibrium.

In the special case of constant entropy and volume, the stability criterion is

$$\Delta E > 0 \quad . \qquad (4.9)$$

That is, the internal energy is a minimum at a stable equilibrium.

For the important special case of an isolated system, $dE + p\,dV = 0$, so the stability criterion becomes

$$\Delta S < 0 \quad . \qquad (4.10)$$

That is, entropy is a maximum at stable equilibrium.

If now the entropy density s is considered to be a function of a set of independent thermodynamic variables (e.g., u, v, and c_γ), then Δs may be written as

$$\Delta s = \delta s + \frac{1}{2}\,\delta^2 s + \ldots \qquad (4.11)$$

where δs is of first order in the deviations from equilibrium, $\delta^2 s$ is of second order, etc. Then, in order for the system to be in equilibrium (entropy a maximum), $\delta s = 0$ and the stability criterion for infinitesimal perturbations is

$$\delta^2 S = \int_V \rho\, \delta^2 s\, dV < 0 \quad . \tag{4.12}$$

To find an explicit form for $\delta^2 s$, consider the independent variables u, v, and c_γ. Then, since they are independent,

$$\delta^2 u = \delta^2 v = \delta^2 c_\gamma = 0 \quad . \tag{4.13}$$

Starting from the Gibbs-Duhem relation, eqn. (2.60), gives

$$\delta^2 s = \frac{1}{T}\{\delta p \delta v - \sum_{\gamma=1}^{N} \delta\mu_\gamma \delta c_\gamma - \delta T \delta s\} \quad . \tag{4.14}$$

Writing $\delta\mu_\gamma$ in terms of T, p, and c_γ and manipulating (cf. Glansdorff and Prigogine [1] p. 25) yields

$$\delta^2 s = -\frac{1}{T}\{\frac{c_v}{T}(\delta T)^2 + \frac{\rho}{\chi}(\delta v)^2 + \sum_{\gamma,\gamma'=1}^{N} \mu_{\gamma\gamma'}\delta c_\gamma \delta c_{\gamma'}\} \tag{4.15}$$

where

$$c_v = (\frac{\partial u}{\partial T})_{v,c_\gamma} \quad ; \quad \chi = -\frac{1}{v}(\frac{\partial v}{\partial p})_{T,c_\gamma} \quad ; \quad \mu_{\gamma\gamma'} = (\frac{\partial \mu_\gamma}{\partial c_{\gamma'}})_{T,p,c_\gamma} \quad . \tag{4.16}$$

Thus, the stability criterion for an isolated system is

$$\int \frac{\rho}{T}\{\frac{c_v}{T}(\delta T)^2 + \frac{\rho}{\chi}(\delta v)^2 + \sum_{\gamma,\gamma'=1}^{N} \mu_{\gamma\gamma'}\delta c_\gamma \delta c_{\gamma'}\}dV > 0 \quad . \tag{4.17}$$

Since the perturbations δT, δv, and δc_γ are arbitrary, this implies

$$c_v > 0 \quad \text{(thermal stability),}$$

$$\chi > 0 \quad \text{(mechanical stability),}$$

and

$$\sum_{\gamma,\gamma'=1}^{N} \mu_{\gamma\gamma'} \delta c_\gamma \delta c_{\gamma'} > 0 \text{ (stability with respect to diffusion and chemical reaction).} \qquad (4.18)$$

These are nearly obvious conditions for a system's stability. Having $c_v > 0$ merely means that removing energy must decrease the temperature; $\chi > 0$ just means that the compressibility must resist any volume change, not aid it. The final equation above is also easily interpreted although it is not as familiar as the other two.

This classical theory of stable equilibrium works well in systems for which a thermodynamic potential can be defined. This is quite restrictive since it is known that in an experimental situation, stability is controlled by the boundary conditions. For example, in a solid, thermal equilibrium is established by a boundary condition either of a uniform temperature or of no heat flow on the surface. This is of even more importance in considering non-equilibrium steady state situations. Hence, it is desirable to formulate the stability criterion in terms of the balance equations and boundary conditions.

First, the equilibrium stability theory is reformulated. For the total system in a volume V, bounded by a surface Σ,

$$\frac{\partial S}{\partial t} = P[S] - \phi[S] \tag{4.19}$$

where
$$P[S] = \int_V \sigma_s dV = \int_V \sum_i X_i\, Y_i\, dV \tag{4.20}$$

is a source term and

$$\phi[S] = \int_\Sigma \vec{J}_s \cdot d\vec{\Sigma} \tag{4.21}$$

is an entropy flow term.

Writing

$$S = S_{eq} + \delta S + \frac{1}{2}\,\delta^2 S + \ldots \tag{4.22}$$

where δS is of first order in deviations from equilibrium, $\delta^2 S$ is of second order, etc., gives

$$\frac{\partial S}{\partial t} = \frac{\partial(\delta S)}{\partial t} + \frac{1}{2}\,\frac{\partial(\delta^2 S)}{\partial t} \quad , \tag{4.23}$$

valid to second order.

Next, split $\{P[S] - \phi[S]\}$ into terms of first and second order. First, note that, since $X_i = Y_i = 0$ in equilibrium, $P[S]$ is of second order. Then, write

$$\phi[S] = \phi^{(1)}[S] + \phi^{(2)}[S] \quad ,$$

where $\phi^{(1)}[S]$ is of first order and $\phi^{(2)}[S]$ is of second order. It is possible to find explicit expressions for $\phi^{(1)}[S]$ and $\phi^{(2)}[S]$ from the equation

$$\phi[S] = \int_{\Sigma} \{\rho s\vec{v} + \frac{\vec{J}_q}{T} - \sum_{\gamma=1}^{N} \frac{\mu_\gamma \vec{J}_\gamma^{diff}}{T}\} \cdot d\vec{\Sigma} \quad . \tag{4.24}$$

Since $\vec{v} = 0$ on Σ for a closed system, and if T^{-1} and $\frac{\mu_\gamma}{T}$ are written in terms of their equilibrium values as

$$T^{-1} = T_{eq}^{-1} + \Delta T^{-1}$$

and

$$\mu_\gamma T^{-1} = (\mu_\gamma T^{-1})_{eq} + \Delta(\mu_\gamma T^{-1}) \quad , \tag{4.25}$$

then $$\phi^{(1)}[S] = \int_{\Sigma} \{\frac{\vec{J}_q}{T_{eq}} - \sum_{\gamma=1}^{N} (\frac{\mu_\gamma}{T})_{eq} \vec{J}_\gamma^{diff}\} \cdot d\vec{\Sigma} \tag{4.26}$$

and $$\phi^{(2)}[S] = \int_{\Sigma} \vec{J}_q \Delta T^{-1} - \sum_{\gamma=1}^{N} \Delta(\frac{\mu_\gamma}{T}) \vec{J}_\gamma^{diff}\} \cdot d\vec{\Sigma} \quad . \tag{4.27}$$

Equating the terms of the same order gives

$$\frac{\partial(\delta S)}{\partial t} = - \phi^{(1)}[S]$$

and

$$\frac{1}{2} \frac{\partial(\delta^2 S)}{\partial t} = P[S] - \phi^{(2)}[S] \quad . \tag{4.28}$$

These equations are valid in this form only for systems at rest (no velocity fluctuations). In the presence of velocity fluctuations, $\frac{\partial(\delta S)}{\partial t}$ and $\phi^{(1)}[S]$ contain additional terms of second order.

Starting in equilibrium at $t = 0$,

$$\delta S = \int_{0}^{t} \phi^{(1)}[S]\, dt \quad . \tag{4.29}$$

Note that, if the system is isolated, $\phi^{(1)}[S] = 0$ and the equilibrium condition, $\delta S = 0$, is recovered. For a non-isolated system, it is seen that eqn. (4.29) is a generalized equilibrium condition which implies that a small change in the system's entropy is compensated by a flow through the surface.

For the second order equation, an assumption is made about $\phi^{(2)}[S]$. It is assumed that it is possible to maintain the boundary conditions on the average. This does not necessarily mean that there are no fluctuations on the surface but rather, for example,

$$\int_0^t \Delta(T^{-1})dt = 0 \qquad (4.30)$$

for $t >> \tau$, where τ is some time scale for fluctuations. Since the stability of interest is for time scales greater than that of the fluctuations (these fluctuations may be regarded as the initial perturbations), it is permissible to set $\phi^{(2)}[S] = 0$. Therefore,

$$\frac{1}{2}\frac{\partial(\delta^2 S)}{\partial t} = P[S] \geq 0 \quad . \qquad (4.31)$$

This then gives the condition for stable equilibrium

$$\int_e^f P[S]dt \leq 0 \qquad (4.32)$$

where e is the initial equilibrium state and f is the final state. The criterion may be rewritten in terms of $\delta^2 S$ since

$$\frac{1}{2}\delta^2 S = \int_0^t P[S]\,dt = \Delta S < 0 \quad . \qquad (4.33)$$

Locally, the stability criterion is

$$\delta^2 s < 0 \tag{4.34}$$

which immediately implies that $c_v > 0$, $\chi > 0$, and $\sum_{\gamma,\gamma'=1}^{N} \mu_{\gamma\gamma'} \, \delta c_\gamma \, \delta c_{\gamma'} > 0$ as before.

Thus, even for the more general situations where equilibrium is due to boundary conditions and no thermodynamic potential exists, the same necessary and sufficient conditions for stability are derived as previously. Of course, the above conditions, holding only at equilibrium, cannot ensure that the equilibrium is stable agains finite perturbations.

§4.3 Stability of Non-Equilibrium States

Next, the above theory will be generalized to include non-equilibrium processes. Only linear stability (small perturbations) will be considered. However, this includes hydrodynamic stability problems, such as the onset of convection or of turbulence, as well as purely dissipative processes such as chemical reactions or thermodiffusion.

For the stability of non-equilibrium states, both the assumption of local equilibrium and the concept of Liapounoff functions are used. If the function $y^2(t)$ is chosen as a Liapounoff function (where $y(t)$ is defined by eqn. (4.2)), then a sufficient condition for stability is $\frac{\partial y^2(t)}{\partial t} \leq 0$. If the independent variables are u, v, and c_γ, then the criterion for asymptotic stability becomes

$$\delta u\, \delta\dot{u} + \delta v\, \delta\dot{v} + \sum_{\gamma=1}^{N} \delta c_\gamma\, \delta\dot{c}_\gamma < 0 \quad . \qquad (4.35)$$

Note that if the quantities δu, δv, and δc_γ are the deviations from a steady state, then $\dot{u} = \delta\dot{u}$, $\dot{v} = \delta\dot{v}$, and $\dot{c}_\gamma = \delta\dot{c}_\gamma$ so the conditions for stability of a steady state are

$$\dot{u}\,\delta u < 0 \ ,$$

$$\dot{v}\,\delta v < 0 \ ,$$

and

$$\dot{c}_\gamma\,\delta c_\gamma < 0 \quad . \qquad (4.36)$$

However, since the quantities appearing in these inequalities do not appear directly in either the balance equations or the boundary conditions, another Liapounoff function is sought.

It is at this point that the assumption of local equilibrium is used. It is also assumed (in agreement with experience) that the local equilibrium is stable. In a nonequilibrium state in local equilibrium, with temperature distribution $T(\vec{r})$, the thermal stability condition would be that

$$c_v[T(\vec{r})] > 0 \quad . \qquad (4.37)$$

Thus, for stable local equilibrium, $\delta^2 s < 0$ everywhere. Therefore, $(-\,\delta^2 s)$ may be used as a Liapounoff function, giving as a sufficient condition for local stability,

$$\frac{\partial}{\partial t}\,(\delta^2 s) \geqslant 0 \quad . \qquad (4.38)$$

Integrating over the entire system gives

$$\delta^2 S = \int \rho\delta^2 s \, dV < 0 \qquad (4.39)$$

with the stability criterion

$$\frac{\partial}{\partial t} (\delta^2 S) \geq 0 \quad . \qquad (4.40)$$

Note that this criterion is sufficient only for infinitesimal perturbations. Schlögl [16], using a statistical definition of entropy (essentially at local equilibrium), has shown that, for finite perturbations, $-\delta_{NL} S \geq 0$ and thus that a sufficient condition for stability is

$$\frac{\partial}{\partial t} (\delta_{NL} S) \geq 0 \qquad (4.41)$$

where $\delta_{NL} S$ is the part of ΔS which is non-linear in the deviations from equilibrium. However, since the Gibbs formula is a first order condition, its validity alone is not enough to ensure that $\delta_{NL} S \leq 0$ and it is necessary to assume that S has the same functional form as in equilibrium to all orders. This is a considerable extension of the idea of local equilibrium. Also, the practical value of this condition is limited by the need to calculate ΔS and $\frac{\partial}{\partial t} (\Delta S)$ exactly. In the limit of infinitesimal perturbations, the stability criterion reduces to that of Glansdorff and Prigogine [1], eqn. (4.34).

Note that the stability criterion given above applies to any non-equilibrium state provided the system is in local equilibrium. However, for time dependent states, stability involves the difference of two separate time-dependent solutions corresponding to the perturbed and unperturbed motions of the system in the parameter space. In steady state situations, stability is concerned only with the deviation from a time-independent initial state. It is this simplified situation which is usually considered. However, in situations where more than one stable steady state exists (e.g., thermal steady state of a temperature dependent resistor - Schlögl [17]), the question of whether a perturbation of one steady state would cause the system to evolve towards another steady state may be of interest.

For an example of a stability criterion, consider an isothermal, linear, passive electrical circuit (cf. Bessenrodt [18]). Then,

$$\rho T \delta s = \delta u - \mu \delta \rho - \vec{E} \cdot \delta \vec{D} - \vec{H} \cdot \delta \vec{B} \quad . \qquad (4.42)$$

If it is assumed that there are no density variations ($\delta\rho = 0$), no inductive elements ($\delta\vec{B} = 0$), and that the system is isothermal ($\delta T = 0$), then,

$$\delta T = 0 = \frac{\partial T}{\partial u}\,\delta u + (\vec{\nabla}_D T) \cdot \delta \vec{D} \qquad (4.43)$$

where $$(\vec{\nabla}_D)_i \equiv \frac{\partial}{\partial D_i} \quad .$$

If it is assumed that the dielectric is such that $\vec{\nabla}_D T \approx 0$, then $\delta u = 0$ and

$$\rho T \delta s = - \vec{E} \cdot \delta \vec{D} \tag{4.44}$$

for an ideal capacitive network. Therefore, the stability criterion is

$$\rho T \delta^2 s = - \delta \vec{E} \cdot \delta \vec{D} = - (\vec{\nabla}_D \vec{E}) : \delta \vec{D}\, \delta \vec{D} < 0 \quad . \tag{4.45}$$

For a parallel plate capacitor, $\vec{E} \parallel \vec{D}$, and hence

$$\frac{\partial E}{\partial D} (\delta D)^2 > 0 \tag{4.46}$$

or

$$\frac{\partial E}{\partial D} > 0 \quad . \tag{4.47}$$

If the capacitor has plate area A, plate separation d, charge Q, volume V = Ad, and potential difference U, then

$$T \delta^2 S = T \int_V \rho \delta^2 s \, dV = - V \frac{\partial E}{\partial D} (\delta D)^2 \tag{4.48}$$

and

$$E = \frac{U}{d} \; ; \; D = \frac{Q}{A} = \frac{Qd}{V} \quad .$$

Thus,

$$\frac{E}{D} = \frac{UA}{Qd} \tag{4.49}$$

so

$$\frac{\partial E}{\partial D} = \frac{\partial U}{\partial Q} \frac{A}{d} \tag{4.50}$$

and

$$\delta D = \frac{\delta Q}{A} \quad . \tag{4.51}$$

Therefore, the stability condition becomes

$$T \delta^2 S = - \frac{VA}{d} \frac{\partial U}{\partial Q} \left(\frac{\delta Q}{A}\right)^2 = - \frac{\partial U}{\partial Q} (\delta Q)^2 < 0 \tag{4.52}$$

or

$$\frac{\partial U}{\partial Q} > 0 \quad . \tag{4.53}$$

Hence, capacitance $\equiv \frac{\partial Q}{\partial U} > 0$ for a stable linear capacitor.

This can be generalized to systems with inductances and generators.

It was noted previously that it is not necessary to use $\delta^2 S$ as the Liapounoff function. Instead, any positive definite quadratic function of $\vec{y}(t)$ could be used. However, Glansdorff and Prigogine [1] argue that $\delta^2 S$ and $\frac{\partial(\delta^2 S)}{\partial t}$ involve quantities directly derivable from the balance equations and boundary conditions and hence $\delta^2 S$ is the easiest quadratic form with which to deal. However, a more convincing reason to use $\delta^2 S$ is seen by considering the connection to the theory of fluctuations. Denoting the probability of realizing a given state, n, by W_n, then from statistical mechanics,

$$S_n = k \ln W_n + \text{constant} \tag{4.54}$$

or

$$W_n \propto e^{\frac{1}{k} S_n} . \tag{4.55}$$

If

$$S_n = S_{eq} + (\delta S)_{eq} + \frac{1}{2} (\delta^2 S)_{eq} + \ldots \tag{4.56}$$

then it is known that $\delta S = 0$ at equilibrium and near equilibrium

$$W_n \propto e^{\frac{1}{2k} \delta^2 S} . \tag{4.57}$$

Thus, it is seen that the condition $\delta^2 S < 0$ means that the state with $\delta^2 S = 0$ (i.e., equilibrium state) is the most probable. This, although it shows directly the connection between stability of equilibrium and the condition $\delta^2 S < 0$, is not sufficient to ensure stability. It is also necessary to have any fluctuations die out (i.e., $\frac{\partial(\delta^2 S)}{\partial t} > 0$).

An example which serves to illustrate several of the points previously covered is that of heat conduction. The equations for the heat current $\vec{J}_q$ and the entropy production σ_s are

$$\vec{J}_q = - \lambda(T)\, \vec{\nabla} T = \overleftrightarrow{L} \cdot \vec{X}$$

where

$$\vec{X} \equiv \vec{\nabla}\left(\frac{1}{T}\right) \quad ,$$

$$\sigma_s = \vec{J}_q \cdot \vec{X} = \frac{\lambda(T)}{T^2} (\vec{\nabla} T) \cdot (\vec{\nabla} T) \quad ,$$

and

$$P = \int_V \sigma_s \, dV = \int_V \frac{\lambda(T)}{T^2} (\vec{\nabla} T) \cdot (\vec{\nabla} T) \, dV \quad . \tag{4.58}$$

Splitting dP into $d_x P$ and $d_y P$ as in eq. (3.101) gives

$$dP = \vec{J}_q \cdot d\vec{X} + d\vec{J}_q \cdot \vec{X} = d_x P + d_y P \quad . \tag{4.59}$$

Then,

$$\frac{d_x P}{dt} = \int_V \vec{J}_q \cdot \frac{\partial \vec{X}}{\partial t} \, dV \tag{4.60}$$

and the energy balance equation is

$$\rho \frac{\partial u}{\partial t} = - \vec{\nabla} \cdot \vec{J}_q \quad . \tag{4.61}$$

Assuming only dissipative processes are present (i.e., no convection), define

$$\psi \equiv \rho \frac{\partial T^{-1}}{\partial t} \frac{\partial u}{\partial t} = - \rho \frac{c_v}{T^2} \left(\frac{\partial T}{\partial t}\right)^2 \leq 0 \qquad (4.62)$$

since $c_v > 0$. Then,

$$\psi = - \frac{\partial T^{-1}}{\partial t} \vec{\nabla} \cdot \vec{J}_q = \vec{\nabla} \cdot \left(- \vec{J}_q \frac{\partial T^{-1}}{\partial t}\right) + \vec{J}_q \cdot \frac{\partial}{\partial t} (\vec{\nabla} T^{-1}) \leq 0 \qquad (4.63)$$

and integrating gives

$$\int_V \psi \, dV = - \int_V \vec{\nabla} \cdot \left(\vec{J}_q \frac{\partial T^{-1}}{\partial t}\right) dV + \int_V \vec{J}_q \cdot \frac{\partial}{\partial t} (\vec{\nabla} T^{-1}) dV$$

$$= - \int_\Sigma \vec{J}_q \frac{\partial T^{-1}}{\partial t} \cdot d\vec{\Sigma} + \int_V \vec{J}_q \cdot \frac{\partial}{\partial t} (\vec{\nabla} T^{-1}) dV \leq 0 \ . \qquad (4.64)$$

Assuming that the boundary conditions are time-independent gives

$$\int_V \vec{J}_q \cdot \frac{\partial}{\partial t} (\vec{\nabla} T^{-1}) dV = \int_V \vec{J}_q \cdot \frac{\partial \vec{X}}{\partial t} dV \leq 0 \ . \qquad (4.65)$$

Thus, in heat conduction, the thermodynamic forces change in such a way as to lower the rate of entropy production.

Assuming the quasi-linear phenomenological law of eqn. (4.58) gives

$$\int_V \left[\lambda(T) T^2 \vec{\nabla} T^{-1} \cdot \frac{\partial}{\partial t} (\vec{\nabla} T^{-1})\right] dV \leq 0 \ . \qquad (4.66)$$

Let the steady state solution be $T_0(\vec{r})$. That is, let $T_0(\vec{r})$ be the solution of

$$\vec{\nabla} \cdot [\lambda(T_o)\, \vec{\nabla} T_o] = 0 \quad . \qquad (4.67)$$

Expand $\lambda(T)T^2 = \lambda(T_o)T_o^2 + \delta(\lambda T^2) + \ldots \quad . \qquad (4.68)$

Then to first order in the deviations from steady state,

$$\frac{\partial}{\partial t} \frac{1}{2} \int_V \lambda(T_o)T_o^2\, (\vec{\nabla} T^{-1}) \cdot (\vec{\nabla} T^{-1})\, dV \leq 0 \quad . \qquad (4.69)$$

Therefore, if the local potential ϕ is defined such that

$$\phi(T,T_o) = \frac{1}{2} \int_V \lambda(T_o)T_o^2\, (\vec{\nabla} T^{-1}) \cdot (\vec{\nabla} T^{-1})\, dV \quad , \qquad (4.70)$$

then

$$\Delta\phi \equiv \phi(T,T_o) - \phi(T_o,T_o)$$

$$= \frac{1}{2} \int_V \lambda(T_o)T_o^2[(\vec{\nabla} T^{-1})\cdot(\vec{\nabla} T^{-1}) - (\vec{\nabla} T_o^{-1})\cdot(\vec{\nabla} T_o^{-1})]\, dV \geq 0$$

$$(4.71)$$

and $\Delta\phi = 0$ if and only if $T(\vec{r}) = T_o(\vec{r})$ provided the boundary conditions are time-independent and the system is near its steady state.

Thus, $\Delta\phi \geq 0$ and $\frac{\partial(\Delta\phi)}{\partial t} \leq 0$ and so $\Delta\phi$ is a Liapounoff function which shows the steady state is stable.

In a steady state, $\phi(T,T_o)$ is minimum and

$$\phi(T_o,T_o) = \frac{1}{2} P = \frac{1}{2} \frac{d_i S}{dt} \quad , \qquad (4.72)$$

so this can be regarded as a generalization of the minimum entropy production theorem.

§4.4 Stability and Fluctuations

As before, stability may be interpreted as a stability against fluctuations. Then, $T_0(\vec{r})$ is the average temperature in steady state and $T(\vec{r})$ can be regarded as a fluctuating temperature, near $T_0(\vec{r})$, analogous to equilibrium fluctuations of local temperature. The situation is illustrated in Figs. (4.1) and (4.2).

The idea of a fluctuating temperature is most easily understood in terms of Einstein's fluctuation formula. For fluctuations about equilibrium, to first non-vanishing order, the probability W of a state is

$$W \alpha\, e^{\frac{1}{k}\Delta S} = e^{\frac{1}{2k}\int_V \rho\delta^2 s\, dV} . \qquad (4.73)$$

Hence

$$W \alpha\, e^{-\frac{1}{2k}\int_V \frac{c_v(\delta T)^2}{T_0^2}\, dV} \qquad (4.74)$$

gives the probability of a temperature distribution $T(\vec{r}) \equiv T_0 + \delta T(\vec{r})$ differing from the average equilibrium temperature T_0. The Einstein formula can be generalized to steady state situations as

$$W \alpha\, e^{-\frac{1}{2k}\int_V \frac{c_v(\delta T)^2}{T_0^2(\vec{r})}\, dV} \qquad (4.75)$$

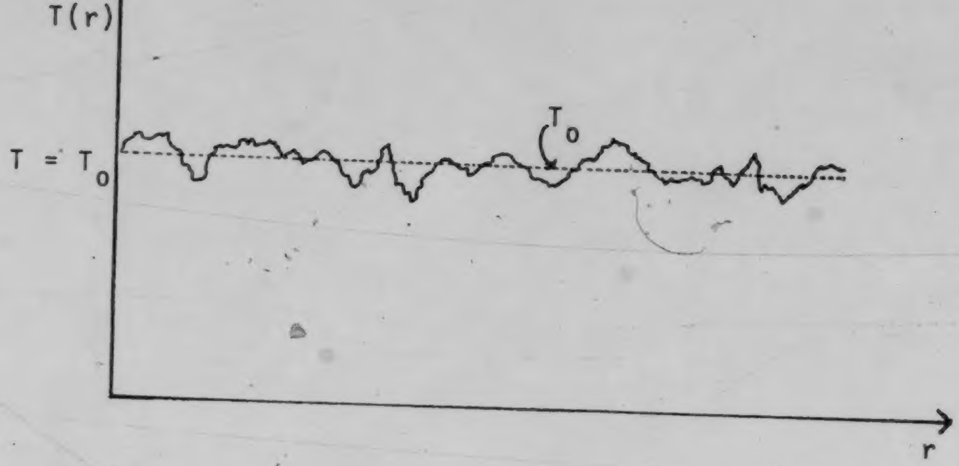


Fig. 4.1 Temperature fluctuations about equilibrium

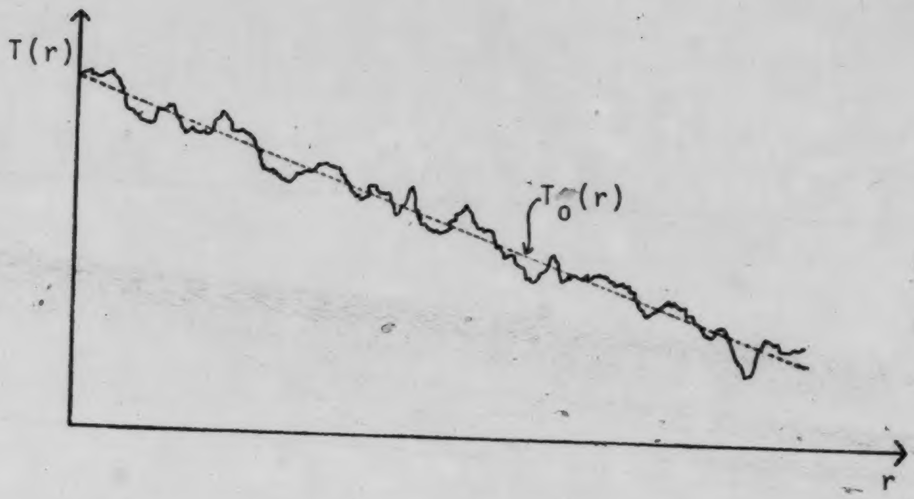


Fig. 4.2 Temperature fluctuations about steady state

where $T_0(\vec{r})$ is the steady state temperature distribution.

However, before this simple interpretation of fluctuations can be accepted, there are some basic questions about the nature of fluctuations in a thermodynamic system. First of all, it is realized that classical equilibrium thermodynamics considers all parameters to be continuous variables which cannot fluctuate. This is basically due to the fact that matter is assumed to be continuous, ignoring its microscopic structure. In order to include this structure in the theory, equilibrium statistical mechanics is needed. This theory admits fluctuations as deviations of statistical quantities from their average values. This may be interpreted as meaning either that the various members of a statistical ensemble have different values of that parameter or that the value of that parameter for a single system fluctuates in time. In order to see which quantities can fluctuate, care must be taken to separate the various ensembles. In the microcanonical ensemble, the mechanical quantities N, V, and E (the number of particles, the volume, and the energy) are fixed so there can be no fluctuations of these extensive parameters. In a canonical ensemble the energy is no longer fixed since the system is not isolated. In such an ensemble, temperature is introduced as a Lagrange multiplier, so it is meaningless to speak of fluctuations in temperature. It is also meaningless to talk, for example, of temperature fluctuations in a microcanonical ensemble since the very concept of temperature is ambiguous in such an ensemble. Indeed, if a "small" thermometer is inserted into an isolated system, "temperature fluctuations" would apparently be observed. However, by inserting the thermometer, energy exchange is allowed and the energy is indefinite after the measurement so the ensemble is no longer microcanonical.

In general, the ensembles considered can only allow fluctuations in the extensive variables (e.g., E, V, N in a grand canonical ensemble). These parameters are mechanically defined and so are unambiguous, regardless of the presence of fluctuations and independently of the size of the system. The thermodynamic intensive variables (e.g., T, p, μ) are then introduced. These are defined unambiguously only for ensembles in which they are constant, hence temperature is ambiguous for a microcanonical ensemble but not for a canonical ensemble where it is constant. This constancy is a reflection of the fact that, in equilibrium, these intensive parameters are defined to be those of the reservoir with which the system is in equilibrium. Thus, these parameters are non-fluctuating quantities by definition. This is also seen in the fact that these parameters enter the theory as Lagrange multipliers and not as statistical quantities.

The apparent conflict with the intuitive feeling that, if particle and energy densities are fluctuating, the temperature cannot be constant is explained as follows.

In equilibrium, the constant intensive parameters refer to the system as a whole. By their very nature, they exclude local variations or fluctuations. However, these intensive parameters might be regarded as averages of some local parameter over the system. It is possible to generalize the concept of temperature to include local fluctuations.

Since the local mechanical densities (e.g., u, ρ, c_γ) are defined in any system, irrespective of fluctuations, a local temperature $T(\vec{r},t)$ may be defined for an equilibrium system as the local function which has exactly the same functional dependence upon the local mechanical densities as the equilibrium temperature T_o has upon the average mechanical densities of the entire system. This "local temperature" will then fluctuate with the local values of the mechanical densities. The reason that this is no longer a constant temperature, as in a canonical ensemble, is that if a small portion of the system is considered, the presence of significant energy and mass density variations in such a region precludes it from being in equilibrium with any reservoir and hence it cannot meaningfully be described as having the temperature of some ideal reservoir.

It should be noted that this local temperature no longer has a close relationship with the Lagrange multiplier T_o which describes the system as a whole. Also, it is obvious that, in order for the local temperature to be a physically meaningful and measurable quantity, its fluctuations must be much smaller than its value. This is also reasonable since, in fixing the functional dependence of $T(\vec{r},t)$, fluctuations were ignored since there were none in T_o.

Also, as is clear from this approach to the local temperature, it is quite unnecessary to have such a function, except as a salve to intuition. What are really of interest are fluctuations in mechanical densities and these fluctuations are cast in terms of a fluctuating temperature only by assuming the functional dependence of $T(\vec{r},t)$ on these mechanical densities.

This is also clear from the approach due to Landau and Lifshitz [19]. There, the expression for δT is

$$\delta T = \frac{\partial T}{\partial u}\bigg|_v \delta u + \frac{\partial T}{\partial v}\bigg|_u \delta v \qquad (4.76)$$

where the partial derivatives are determined from the equilibrium functional dependence of T_o upon (E/V) and (V/N). Thus, the results referring to temperature fluctuations (e.g., eqn. (4.14)) are not in conflict with the idea of intensive thermodynamic parameters as constants. Rather, this is a convenient way of expressing results which originally referred only to fluctuations in u and v. Indeed in Landau and Lifshitz [19], p. 351, the original expression for probability is

$$W \alpha e^{-(\Delta E - T_0\Delta S + P_0\Delta V)/kT_0} \tag{4.77}$$

where P_0 and T_0 are constants, and it is only through a formal procedure, as above, that this becomes

$$W \alpha e^{-(\frac{c_v(\Delta T)^2}{2kT^2} - \frac{1}{2kT}(\frac{\partial P}{\partial V})_T (\Delta V)^2)} \quad . \tag{4.78}$$

CHAPTER V

CHEMICAL REACTIONS

§5.1 Equations of Motion and Stability

Chemically reacting systems will now be considered as examples of systems which may exhibit steady states both near and far from equilibrium as well as other quasi-stable states (e.g., sustained oscillations which may be sensitive to perturbations).

Consider a system of chemical reactants with mole numbers n_γ ($\gamma=1,2,\ldots,c$). If there is one chemical reaction, then

$$dn_\gamma = \nu_\gamma d\,\xi \tag{5.1}$$

where ν_γ is the stoichiometric coefficient for the γ^{th} component and ξ is a parameter describing the advancement of the reaction.

If there are "r" simultaneous reactions then

$$dn_\gamma = \sum_{\delta=1}^{r} \nu_{\gamma\delta}\ d\,\xi_\delta \tag{5.2}$$

where $\nu_{\gamma\delta}$ is the stoichiometric coefficient of the γ^{th} reactant in the δ^{th} reaction and ξ_δ is the advancement of the δ^{th} reaction.

If the velocity v_δ of the δ^{th} reaction is defined as

$$v_\delta = \frac{d\,\xi_\delta}{dt}\quad , \tag{5.3}$$

then

$$\frac{dn_\gamma}{dt} = \sum_{\delta=1}^{r} \nu_{\gamma\delta} v_\delta \quad . \tag{5.4}$$

To clarify the notation, consider the simultaneous reactions:

$$\begin{aligned} 2C + O_2 &\rightarrow 2CO \qquad (\delta=1) \\ C + O_2 &\rightarrow CO_2 \qquad (\delta=2) \quad . \end{aligned} \tag{5.5}$$

Then,

$$dn_c = -2\, d\xi_1 - d\xi_2 \quad ,$$

$$dn_{O_2} = -\, d\xi_1 - d\xi_2 \quad ,$$

$$dn_{CO} = 2\, d\xi_1 \quad ,$$

and

$$dn_{CO_2} = d\xi_2 \quad . \tag{5.6}$$

If the possibility of open chemical systems is included in order to have steady state reactions, the equations become

$$dn_\gamma = d_e n_\gamma + \sum_{\delta=1}^{r} \nu_{\gamma\delta}\, d\xi_\delta \tag{5.7}$$

where $d_e n_\gamma$ is the external supply of reactant γ .

Recalling now the Gibbs-Duhem relation gives

$$dS = \frac{dQ}{T} + \sum_{\delta=1}^{r} \frac{A_\delta\, d\xi_\delta}{T} \tag{5.8}$$

where the affinity of the δ^{th} reaction, A_δ , is defined as

$$A_\delta = \sum_{\gamma=1}^{c} \nu_{\gamma\delta}\, \mu_\gamma \quad . \tag{5.9}$$

As usual, the internal part of the entropy change is considered, yielding the entropy production P as

$$P \equiv \frac{d_i S}{dt} = \frac{1}{T} \sum_{\delta=1}^{r} A_\delta v_\delta \geq 0 \quad . \qquad (5.10)$$

Also, recall the Gibbs-Duhem relation for chemical stability:

$$\sum_{\gamma,\gamma'=1}^{c} \frac{\partial \mu_\gamma}{\partial n_{\gamma'}} \frac{dn_\gamma}{dt} \frac{dn_{\gamma'}}{dt} = - \sum_{\delta=1}^{r} v_\delta \frac{dA_\delta}{dt} > 0$$

or

$$\sum_{\delta=1}^{r} v_\delta \frac{dA_\delta}{dt} < 0 \quad . \qquad (5.11)$$

If the affinities are regarded as the generalized forces and the reaction velocities as the currents, then

$$T\, d_X P = \sum_{\delta=1}^{r} v_\delta\, dA_\delta < 0 \quad . \qquad (5.12)$$

In a steady state, the forces and fluxes are constant so $d_X P = 0$ which implies that

$$\sum_{\delta=1}^{r} v_\delta\, dA_\delta = 0 \quad . \qquad (5.13)$$

§5.2 Examples

The following example serves to illustrate some of these results. Consider three reactions:

$$\begin{aligned} &(1)\ A \rightleftarrows X \\ &(2)\ X \rightleftarrows B \\ &(3)\ X \rightleftarrows M \quad . \end{aligned} \qquad (5.14)$$

The affinities are

$$A_1 = \mu_A - \mu_X$$

$$A_2 = \mu_X - \mu_B$$

$$A_3 = \mu_X - \mu_M \quad . \qquad (5.15)$$

To get a steady state situation, keep n_A and n_B fixed by supplying them externally. This means that μ_A and μ_B are fixed. Then,

$$A_1 + A_2 = \mu_A - \mu_B = \text{constant} \qquad (5.16)$$

so

$$dA_1 + dA_2 = 0 \quad . \qquad (5.17)$$

Since this is a steady state,

$$T\, d_x P = 0 = \sum_{\delta=1}^{3} v_\delta \, dA_\delta \qquad (5.18)$$

or

$$(v_1 - v_2)dA_1 + v_3 \, dA_3 = 0 \qquad (5.19)$$

for any dA_1 and dA_3. Therefore, $v_1 = v_2$ and $v_3 = 0$.

These results could, of course, have been easily deduced by inspection of the reaction, subject to the constraints that the system be in a steady state and that the reactant "M" is not externally supplied.

In the case of steady states near to equilibrium, it may be assumed that the equations are linear in the deviations from equilibrium. For example, in the linear range,

$$v_\delta = \sum_{\delta'=1}^{c} \alpha_{\delta\delta'} \, d\,\xi_{\delta'}$$

and

$$A_\delta = \sum_{\delta'=1}^{c} \beta_{\delta\delta'} \, d\,\xi_{\delta'} \qquad (5.20)$$

which may be combined to yield linear phenomenological laws of the form

$$v_\delta = \sum_{\delta'=1}^{c} L_{\delta\delta'} \, A_{\delta'} \quad . \qquad (5.21)$$

Note that $v_\delta = A_\delta = 0$ in equilibrium.

These linear laws would be expected to obey Onsager relations (within the limitations of §3.2).

Due, however, to the complex form of the chemical potentials and affinities, it would be desirable to express the reaction rates directly in terms of the reactant densities n_δ . For example, for an ideal gas,

$$\mu_\delta = -RT \ln \left[\frac{(2\pi mkT)^{3/2}}{h^3} \frac{V}{N_\delta}\right] \qquad (5.22)$$

and, in the previous example,

$$A_1 = \mu_A - \mu_X = RT \ln \left[\left(\frac{m_X}{m_A}\right)^{3/2} \frac{n_A}{n_X}\right] \quad . \qquad (5.23)$$

In the light of this complex dependence of the affinities upon the observed quantities n_δ, it is usually easier to guess the dependence of the reaction rates upon the concentrations, rather than upon the affinities themselves.

For example, considering the transformations of A particles into B particles through collisions with another B particle in a dilute gas, it is reasonable to assume that the rate at which the A particles transform into B particles is proportional to the concentration of A particles and to the concentration of B particles. Also, if there is a reverse reaction with two B's colliding to yield an A and a B, its rate would be proportional to the square of the concentration of B particles. That is, for the reaction

$$A + B \rightleftarrows 2B \tag{5.24}$$

it would be assumed that

$$\frac{dn_A}{dt} = \kappa_1 n_A n_B - \kappa_1' n_B^2 \quad . \tag{5.25}$$

As an example, consider the stability and steady state properties of a slightly more complicated model of a chemical reaction due to Schlögl [20]. The relevant reactions are

$$A + 2X \rightarrow 3X$$

and

$$B + X \rightarrow C \quad . \tag{5.26}$$

The concentrations of A, B, and C are held constant by an external supply. Only the amount of X can vary and the steady state occurs when $\frac{dn_X}{dt} = 0$. The rates r_1, at which A particles disappear, and r_2, at which B particles disappear, are assumed to be

$$r_1 = \kappa_1 n_A n_X^2 - \kappa_1' n_X^3$$

and

$$r_2 = \kappa_2 n_B n_X - \kappa_2' n_c \quad . \tag{5.27}$$

With an appropriate choice of units, $\kappa_1' = 1$ and $\kappa_1 n_A = 3$. The net rate of change of n_X is then

$$\frac{dn_X}{dt} = r_1 - r_2 = \phi(n_X) = - n_X^3 + 3n_X^2 - \beta n_X + \gamma \qquad (5.28)$$

where $\beta = \kappa_2 n_B$ and $\gamma = \kappa_2' n_c$.

The steady states of the system are the solutions of the equation

$$\gamma = n_X^3 - 3n_X^2 + \beta n_X \quad . \qquad (5.29)$$

A rough plot of γ as a function of n_X for different values of β is given in Fig. 5.1.

For $\beta \geqslant 3$, there is only one steady state value of n_X for any γ. For $\beta < 3$, there are three separate steady state values $n_X^{(1)}$, $n_X^{(2)}$, and $n_X^{(3)}$, for $\gamma_1 < \gamma < \gamma_2$. For γ outside this range, there is just one steady state.

By considering the sign of $\frac{dn_X}{dt}$ near each of the steady states, it is easily seen that $n_X^{(1)}$ and $n_X^{(3)}$ correspond to stable steady states, while $n_X^{(2)}$ corresponds to an unstable steady state.

It has been pointed out by Schlögl [20] that it is possible to have both stable steady states existing simultaneously in a single system, if diffusion is included in the evolution equations. Under suitable conditions, there can be two coexisting "phases" in a vessel, separated by a boundary layer thin with respect to the dimensions of the system.

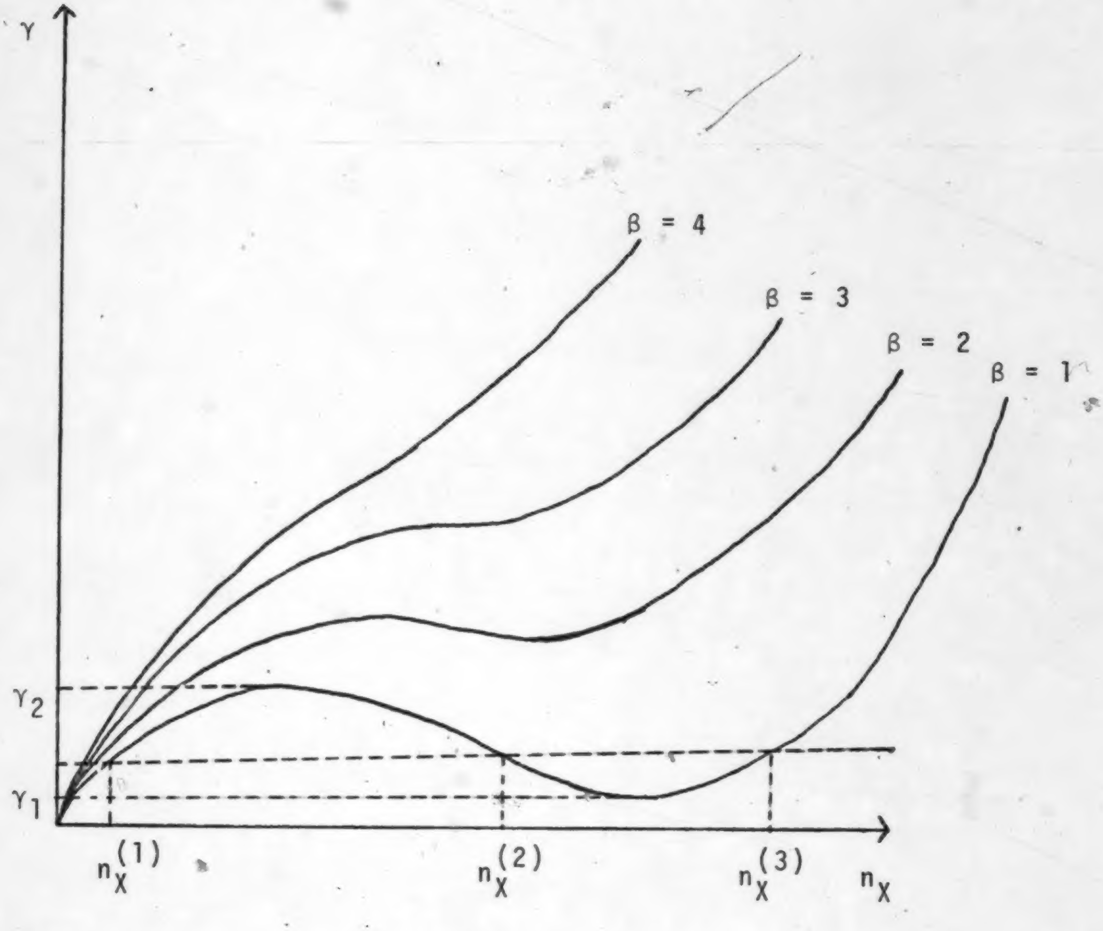


Fig. 5.1 γ versus n_X in a reacting system

For some chemical systems, it is also possible to describe the motion slightly away from steady state. In the following model, a steady state is found which is not asymptotically stable and sustained oscillations of the reactant concentrations result.

Consider the following system of reactions:

$$\begin{aligned} A + X &\rightleftarrows 2X \quad (\text{rates } \kappa_1, \kappa_1') \\ X + Y &\rightleftarrows 2Y \quad (\text{rates } \kappa_2, \kappa_2') \\ Y &\rightleftarrows E \quad (\text{rates } \kappa_3, \kappa_3') \quad . \end{aligned} \tag{5.30}$$

The concentrations of the initial reactant A and the final product E are externally maintained at constant values. Only the concentrations of the intermediate reactants X and Y are variable. The equations for n_X and n_Y are

$$\frac{dn_X}{dt} = \kappa_1 n_A n_X - \kappa_1' n_X^2 - \kappa_2 n_X n_Y + \kappa_2' n_y^2$$

and

$$\frac{dn_Y}{dt} = \kappa_2 n_X n_Y - \kappa_2' n_Y^2 - \kappa_3 n_Y + \kappa_3' n_E \tag{5.31}$$

where the primes denote the reverse reaction rates.

Now, assume that the concentrations are maintained at values for which the affinities become so large that the reactions proceed in the forward direction only. Then, ignoring reverse reactions,

$$\frac{dn_X}{dt} = \kappa_1 n_A n_X - \kappa_2 n_X n_Y$$

and

$$\frac{dn_Y}{dt} = \kappa_2 n_X n_Y - \kappa_3 n_Y \quad . \tag{5.32}$$

This system, incidentally, is isomorphic to the Lotka-Volterra model of prey-predator interactions.

Eqns. (5.32) have one steady state solution (apart from the trivial one $n_X = n_Y = 0$). That solution is

$$n_X^o = \frac{\kappa_3}{\kappa_2} \quad \text{and} \quad n_Y^o = \frac{\kappa_1}{\kappa_2} n_A \quad . \tag{5.33}$$

To analyse the stability of this steady state, a normal mode analysis is used for small perturbations. Assume a solution of the form

$$\begin{aligned} n_X(t) &= n_X^o + \delta n_X \, e^{\omega t} \\ n_Y(t) &= n_Y^o + \delta n_Y \, e^{\omega t} \end{aligned} \tag{5.34}$$

where $|\delta n_X| << |n_X^o|$, $|\delta n_Y| << |n_Y^o|$, and ω may be complex. Then, to first order in the perturbations,

$$\begin{aligned} \omega \, \delta n_X + \kappa_3 \, \delta n_Y &= 0 \\ \text{and} \qquad - \kappa_1 \, n_A \, \delta n_X + \omega \, \delta n_Y &= 0 \quad . \end{aligned} \tag{5.35}$$

These equations imply

$$\omega^2 = - \kappa_1 \, \kappa_3 \, n_A \tag{5.36}$$

so ω is purely imaginary and there are undamped periodic fluctuations about the steady state for infinitesimal perturbations from that steady state. This is an example of weak (non-asymptotic) stability with regard to infinitesimal perturbations. For such perturbations, it can be shown (cf. Glansdorff and Prigogine [1] p. 226) that

$$\delta_X P = 0$$

so

$$\frac{\partial(\delta^2 S)}{\partial t} = 0 \qquad (5.37)$$

and hence $\delta^2 S$ is a non-zero constant of the motion.

For finite perturbations, first eliminate dt from the equations, yielding

$$\frac{dn_Y}{dn_X} = -\frac{n_Y(n_X - \frac{\kappa_3}{\kappa_2})}{n_X(n_Y - \frac{\kappa_1}{\kappa_2} n_A)} \quad . \qquad (5.38)$$

Thus, $n_X + n_Y - \frac{\kappa_1}{\kappa_2} n_A \ln n_Y - \frac{\kappa_3}{\kappa_2} \ln n_X = \text{constant}$. (5.39)

This equation represents a family of closed curves, each corresponding to a different value of the constant in eqn. (5.39). Examples of such curves are illustrated in Fig. 5.2.

Thus, the actual motion in $n_X - n_Y$ space is periodic for any perturbation. The period depends upon the particular initial perturbation. Once the system is perturbed, it will never return to the steady state, but will oscillate about it in some non-linear fashion. Also, the closer an "orbit" comes to the origin, the further the system goes from the origin at the opposite end of the orbit. These features reflect the fact that there are no dissipative processes represented in the equations which could act to limit the orbits and return the system to its steady state.

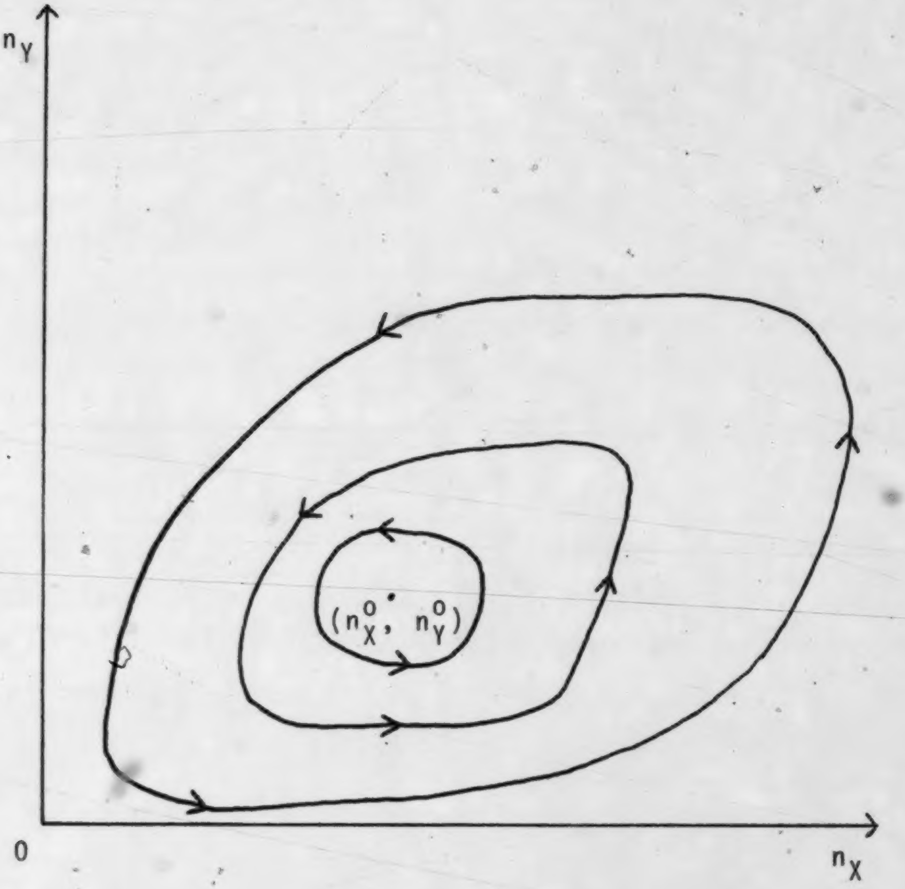

n_Y
(n_X^o, n_Y^o)
0
n_X

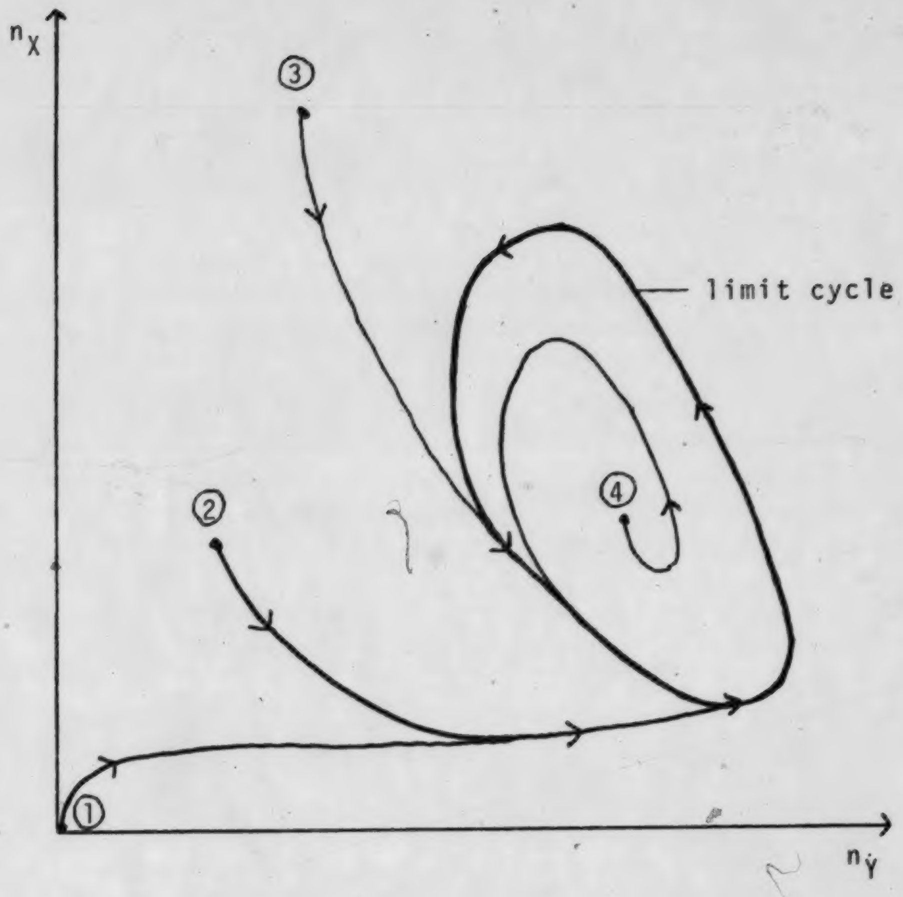
Fig. 5.2 Solutions of eqn. (5.39)

Since this system will oscillate with different frequencies and different orbits, depending upon the initial perturbation, the system would be expected to exhibit highly irreproducible oscillations in a physical system. In addition, even the orbits are not stable in the sense of Liapounoff, since a small perturbation of an orbit results in a new orbit with a different frequency so the perturbed and unperturbed motions will tend to separate. However, the orbits are stable in the more general sense that a perturbation will result in a new orbit, the curve of which will still be close to the curve of the original orbit.

It is possible to have a similar chemical system in which the final state, although it is an oscillatory, non-steady state, is independent of the initial conditions. This means that there is a limit cycle toward which all motions tend as $t \to \infty$, regardless of the size or nature of the initial perturbation. This behavior is illustrated in Fig. 5.3.

In order to get such behavior, a feedback mechanism is needed. Such a mechanism may be provided by an autocatalytic reaction. Consider the reactions

$$\begin{aligned} A &\rightleftarrows X \\ 2X + Y &\rightleftarrows 3X \quad \text{(autocatalytic)} \\ B + X &\rightleftarrows Y + D \\ X &\rightleftarrows E \end{aligned} \qquad (5.40)$$



n_X
n_Y
①
②
③
④
limit cycle

Fig. 5.3 A system above critical affinity

corresponding to the overall reaction

$$A + B \rightleftarrows E + D \quad . \qquad (5.41)$$

This example is due to Glansdorff and Prigogine [1] p. 233.

This is an open system, since the concentrations of A, B, E, and D are held fixed. It is also far from equilibrium, incorporates a feedback mechanism, and has the interesting property that below a certain critical affinity for the overall reaction (eqn. (5.41)), the single non-trivial steady state is a point of asymptotic stability and above the critical affinity, the system is unstable. The system exhibits instability in a manner similar to the Lotka-Volterra model of eqn. (5.32). However, in this case, the orbits are not closed but rather all spiral towards a limit cycle. Thus, as t → ∞ , the system oscillates with a frequency independent of the initial conditions.

CHAPTER VI

STABILITY IN FLUIDS AT REST (BÉNARD PROBLEM)

§6.1 Statement of Problem

In order to gain a better insight into the nature of instabilities, the Bénard problem will be considered. This problem concerns a horizontal layer of fluid of depth d, heated from below. Due to thermal expansion, the warmer fluid at the bottom experiences a bouyant force. This is counteracted by the retarding effect of viscosity. At the critical temperature gradient where the bouyant forces exceed the viscous ones, the system will become unstable and the fluid will begin to convect. The treatment given here will be quite short and for a more detailed discussion of the validity of the various approximations and of the results, see Chandrasekhar [21].

First, write the equations governing the system which is supposed to be so large horizontally that there are no edge effects present. These equations are

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot(\rho\vec{v}) = 0 \quad \text{(continuity)} \quad , \qquad (6.1)$$

$$\rho\left(\frac{\partial\vec{v}}{\partial t} + \vec{v}\cdot\vec{\nabla}\vec{v}\right) + \vec{\nabla}\cdot\overleftrightarrow{P} = \rho\vec{F} \quad \text{(momentum)} \quad , \qquad (6.2)$$

and

$$\left(\frac{\partial(c_v T)}{\partial t} + \vec{v}\cdot\vec{\nabla}(c_v T)\right) = \vec{\nabla}\cdot(\lambda\vec{\nabla}T) - p\vec{\nabla}\cdot\vec{v} - \phi$$

$$\text{(energy balance)} \qquad (6.3)$$

where

$$\phi = -\overleftrightarrow{\Pi} : \vec{\nabla}\vec{v} \qquad (6.4)$$

and the following simplifying assumptions have been made:

(1) there is only one fluid component

(2) isotropic, Newtonian fluid

(3) Fourier's law of heat conduction holds

(4) linear thermal expansion (i.e., $\rho = \rho_0 [1 - \alpha(T-T_0)]$)

(5) specific heat is defined as $c_v = \frac{u}{T}$.

Despite these simplifications, the equations are quite general. However, for many liquids and for moderate temperature ranges, it is possible to simplify these equations further, due to the smallness of the thermal expansion. Then, ρ is treated as a constant, except in the momentum equation (6.2), where it is important since it is the cause of the instability. Also, η, α, and λ are considered to be constants. Since ρ does not change, the term $p\vec{\nabla} \cdot \vec{v}$ can also be ignored. From order of magnitude estimates, ϕ can be neglected in the energy equation (6.3). With these approximations (Boussinesq approximation), the equations become

$$\vec{\nabla} \cdot \vec{v} = 0 \quad , \tag{6.5}$$

$$\frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot (\vec{\nabla}\vec{v}) + \frac{1}{\rho_0} \vec{\nabla} p = (1 + \frac{\delta\rho}{\rho_0})\vec{F} + \frac{\eta}{\rho_0} \nabla^2 \vec{v} \quad , \tag{6.6}$$

and

$$\frac{\partial T}{\partial t} + \vec{v} \cdot \vec{\nabla} T = \frac{\lambda}{\rho_0 c_v} \nabla^2 T \tag{6.7}$$

where

$$\rho = \rho_0 + \delta\rho \quad .$$

Now, consider the initial condition in which a steady temperature gradient is maintained and the fluid is at rest. Then, the temperature and density as functions of z (the vertical height above the bottom of the fluid) are

$$T = T_0 - \beta z$$

$$\rho = \rho_0 (1 + \alpha\beta z) \qquad (6.8)$$

where β is the adverse temperature gradient. Then, since the equation for pressure is

$$\frac{\partial p}{\partial z} = - g\rho \qquad (6.9)$$

where g is the gravitational acceleration, the pressure is found to be

$$p = p_0 - g\rho_0 \left(z + \frac{1}{2}\alpha\beta z^2\right) \quad . \qquad (6.10)$$

To study the stability of such a steady state, the time development of the state, after it is slightly perturbed, is studied. If $\vec{v}$ is the velocity in the perturbed state, δp is the change in pressure, and if the perturbed temperature distribution is

$$T' = T_0 - \beta z + \theta \qquad (6.11)$$

then, considering only first order terms in the perturbation, the Boussinesq approximate equations for the perturbations are

$$\vec{\nabla} \cdot \vec{v} = 0$$

$$\frac{\partial \vec{v}}{\partial t} = - \vec{\nabla}(\frac{\delta p}{p_o}) + \frac{\mu}{\rho_o} \nabla^2 \vec{v} + g\alpha\theta\hat{z}$$

$$\frac{\partial \theta}{\partial t} = \beta v_z + \frac{\lambda}{\rho_o c_v} \nabla^2 \theta \tag{6.12}$$

where $\hat{z}$ is a unit vector in the z direction. The term $\vec{\nabla}(\frac{\delta p}{p_o})$ can be eliminated by taking the curl of the second of eqns. (6.12). Writing $\vec{\omega} = \vec{\nabla} \times \vec{v}$ and taking the curl again gives

$$\frac{\partial \omega_z}{\partial t} = \nu \nabla^2 \omega_z \tag{6.13}$$

where $\nu \equiv \eta/\rho_o$ is the kinematic viscosity and

$$\frac{\partial(\nabla^2 v_z)}{\partial t} = g\alpha(\frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2}) + \nu \nabla^4 v_z \quad . \tag{6.14}$$

These are the required perturbation equations. The boundary conditions are

$$\theta(z=0) = \theta(z=d) = v_z(z=0) = v_z(z=d) = 0 \quad . \tag{6.15}$$

The remaining boundary conditions depend upon whether the surfaces are rigid or free. For a rigid surface, $v_x = v_y = 0$ and, since $\vec{\nabla} \cdot \vec{v} = 0$, $\frac{\partial v_z}{\partial z} = 0$ at a rigid boundary. For a free surface,

$$\overleftrightarrow{P}_{xz} = \overleftrightarrow{P}_{yz} = 0 \quad . \tag{6.16}$$

This implies

$$\eta\left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right) = \eta\left(\frac{\partial v_y}{\partial z} + \frac{\partial v_z}{\partial y}\right) = 0 \quad . \qquad (6.17)$$

Therefore, since $v_z = 0$ on the surface,

$$\frac{\partial v_x}{\partial z} = \frac{\partial v_y}{\partial z} = 0 \qquad (6.18)$$

and from eqn. (6.5),

$$\frac{\partial^2 v_z}{\partial z^2} = 0 \qquad (6.19)$$

on a free surface.

Similarly, for the vorticity $\vec{\omega}$, the boundary conditions are $\omega_z = 0$ for a rigid surface and $\frac{\partial \omega_z}{\partial z} = 0$ for a free surface.

In order to study the stability of the steady state, the perturbation is analysed into a complete set of normal modes and the time dependence of each mode is studied. Therefore, assume the perturbation has the form

$$v_z = V(z)\, e^{[i(k_x x + k_y y) + pt]}$$

$$\omega_z = W(z)\, e^{[i(k_x x + k_y y) + pt]}$$

$$\theta = \Theta(z)\, e^{[i(k_x x + k_y y) + pt]} \qquad (6.20)$$

where p may be complex. The equations are then

$$p\left(\frac{d^2}{dz^2} - k^2\right)V(z) = -\, g\alpha k^2\, \Theta(z) + \nu\left(\frac{d^2}{dz^2} - k^2\right)^2 V(z)$$

$$p\,\Theta(z) = \beta V(z) + \frac{\lambda}{\rho_o c_v}\left(\frac{d^2}{dz^2} - k^2\right)\Theta(z)$$

$$p\,W(z) = \nu\left(\frac{d^2}{dz^2} - k^2\right) W(z) \quad . \qquad (6.21)$$

Choosing units of length to be (d) and of time to be $(\frac{d^2}{\nu})$ and defining the constants $a \equiv kd$ and $\sigma \equiv \frac{pd^2}{\nu}$, the equations become

$$(D^2 - a^2)(D^2 - a^2 - \sigma)\, V(z) = \left(\frac{g\alpha}{\nu}\, d^2\right)a^2\, \Theta(z)$$

and $$(D^2 - a^2 - \kappa\sigma)\,\Theta(z) = -\left(\frac{\beta\, d^2\, \rho_o\, c_v}{\lambda}\right) V(z) \qquad (6.22)$$

where $$\kappa \equiv \frac{\nu\, \rho_o\, c_v}{\lambda} \qquad (6.23)$$

is the Prandtl number and where $D \equiv \frac{d}{dz}$. Eliminating $\Theta(z)$ gives

$$(D^2 - a^2)(D^2 - a^2 - \sigma)(D^2 - a^2 - \kappa\sigma)V(z) = -\, Ra^2 V(z) \qquad (6.24)$$

where $$R \equiv \frac{g\alpha\beta\rho_o c_v}{\nu\lambda}\, d^4 \qquad (6.25)$$

is the Rayleigh number.

It is possible to show that σ is purely real and hence that if $\sigma > 0$, the perturbation grows and the system is unstable. The condition for marginal stability is $\sigma = 0$. Eqn. (6.24) with $\sigma = 0$, along with appropriate boundary conditions, forms an eigenvalue problem. After finding the lowest eigenvalue of R as a function of a^2, the minimum of R with respect to a^2 must be found. This value, called R_c, is the lowest value of R for which instability will arise.

It is possible to solve this eigenvalue problem exactly, and the result, if both surfaces are free, is that

$$V(z) = A \sin(n\pi z)$$

and

$$R = \frac{(n^2 \pi^2 + a^2)^3}{a^2} \qquad (6.26)$$

where n is an integer. Therefore,

$$R_c = \frac{27}{4} \pi^4 \approx 657 \qquad (6.27)$$

at a critical value of a^2 of

$$a_c^2 = \frac{\pi^2}{2} \quad . \qquad (6.28)$$

These unstable disturbances will first appear with a characteristic wavelength

$$\lambda_c = \frac{2\pi d}{a_c} = 2^{3/2} d \quad . \qquad (6.29)$$

For other boundary conditions the critical Rayleigh number is different (e.g., for rigid boundaries, $R_c \approx 1708$).

Thus, in a given experiment, the value of R is set by fixing the value of β, the temperature gradient, and for $R < R_c$ the system is stable. For $R > R_c$, the system is unstable and spontaneously starts to convect in a cell pattern with cell size determined by the value of d.

These results have been verified within the limits of experimental accuracy. As an example, for rigid boundaries, the experimental value of R_c is 1700 ± 51, in good agreement with the theoretical value.

The above theory of the Bénard instability has the simple interpretation that the instability occurs at the minimum temperature gradient at which a balance can steadily be maintained between the kinetic energy dissipated by viscosity and the internal energy released by the bouyant force (cf. Chandrasekhar [21]).

§6.2 Stability Criterion of Glansdorff and Prigogine

Glansdorff and Prigogine [1] have analysed the Bénard problem in terms of Liapounoff functions rather than normal modes. The stability criterion which they derive is a sufficient condition and hence is less precise than the exact solution by normal mode analysis. Their approach will be briefly outlined. In the spirit of section 4.1, thermodynamic stability is necessary in the sense that entropy fluctuations must regress. This can be expressed as the requirement that

$$\frac{\partial(\varepsilon^2\delta^2 s)}{\partial t} > 0 \qquad (6.30)$$

where ε^2 is some positive definite weighting function. However, this criterion does not apply to systems in which there are convective motions. In this case, velocity fluctuations must also regress. This requirement can be written as

$$\frac{\partial(\tau^2(\delta v)^2)}{\partial t} < 0 \qquad (6.31)$$

where δv is the difference between the fluctuating and steady state velocities and τ^2 is another weighting function. Using the balance equations, these criteria can be written as

$$\int_0^d \{\frac{\lambda}{\rho_0 c_v} <(\vec{\nabla}\theta) \cdot (\vec{\nabla}\theta)> - \beta<\theta v_z>\}dz > 0 \qquad (6.32)$$

and $$\int_0^d \{\nu <(\frac{\partial v_\alpha}{\partial x_\beta})^2> - g\alpha<\theta v_z>\}dz > 0 \qquad (6.33)$$

where < > denotes averaging over the horizontal plane.

The first is a purely thermodynamic criterion, while the second is a hydrodynamic stability criterion analogous to the stability criterion in normal mode analysis. Each of these criteria separately yields a critical Rayleigh number and it is the smaller of the two which gives the desired stability criterion, since both (6.32) and (6.33) must be satisfied for the system to be stable. Note that these criteria are more severe than the normal mode criterion since they require that each decay be monotonic. Also, note that, at the instability determined by the normal mode method, the fluctuations themselves diverge, so $(-\delta^2 s)$ and $(\delta v)^2$ must also diverge. Thus, if the critical Rayleigh number for the normal mode approach is denoted by R_c^{NM}, and those for the entropy and velocity regression criteria (eqns. (6.30) and (6.31)) are denoted by R_c^s and R_c^v respectively, then both R_c^s and R_c^v are less than or equal to R_c^{NM}. They will equal R_c^{NM} if there are no oscillatory solutions of the perturbation equations since then the normal mode criterion implies that all disturbances decay monotonically. This means $\frac{\partial(\delta^2 s)}{\partial t} > 0$ and $\frac{\partial(\delta v)^2}{\partial t} < 0$ and the separate criteria are equivalent.

Glansdorff and Prigogine combine the two separate criteria (6.30) and (6.31) into one by choosing as a Liapounoff function

$$\zeta^2 \equiv \tau^2(\delta v)^2 - \varepsilon^2 \, \delta^2 s \qquad (6.34)$$

where ε^2 and τ^2 are arbitrary positive weighting functions with $\delta\tau^2 = \delta\varepsilon^2 = 0$. This stability criterion is less restrictive than the two taken separately. This is most easily seen by considering a path in $-\varepsilon^2 \, \delta^2 s$, $\tau^2(\delta v)^2$ space which is stable with respect to the criterion

$$\frac{\partial\zeta^2}{\partial t} < 0 \qquad (6.35)$$

but not with respect to the combination of (6.30) and (6.31), as illustrated in Figs. 6.1a, 6.1b, and 6.1c. The same evolution is illustrated in all three graphs. It is obvious from Fig. 6.1a that this path satisfies eqn. (6.35) and so that the system the graph refers to is stable by that criterion, and it is equally obvious from Fig. 6.1c that criterion (6.31) is violated. Thus, a system may evolve in

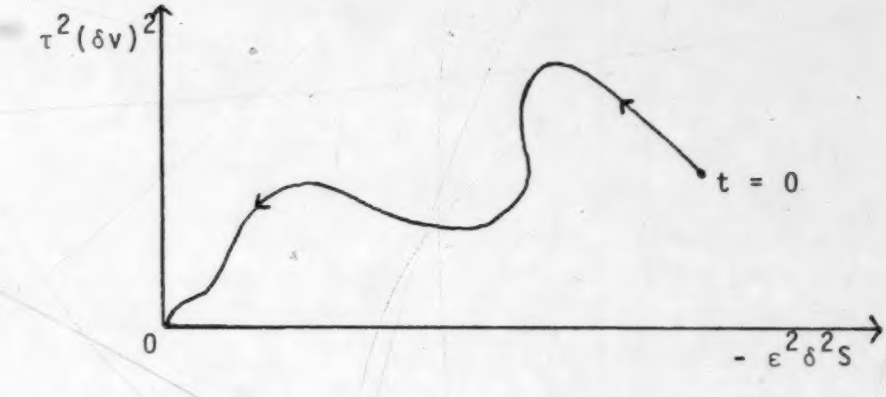


Fig. 6.1a

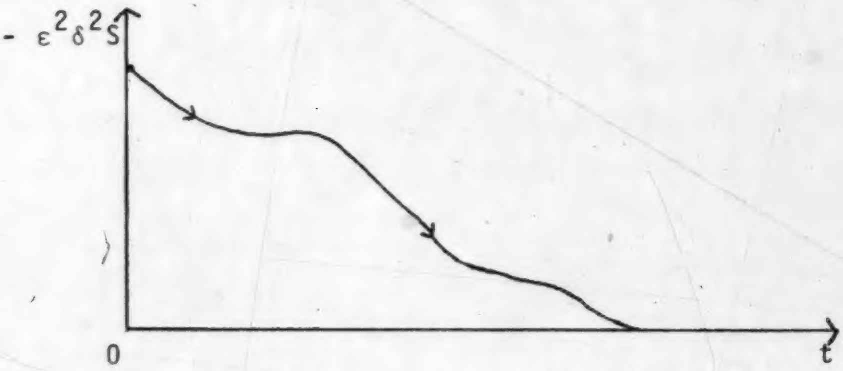


Fig. 6.1b

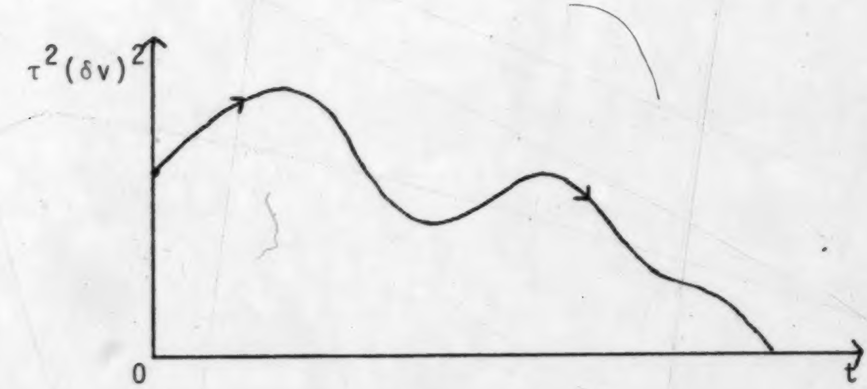


Fig. 6.1c
A possible evolution and its stability

such a manner as to be stable with respect to eqn. (6.35) but not with respect to both of eqns. (6.30) and (6.31).

However, although eqn. (6.35) is a less restrictive criterion, it is still not as general as the normal mode analysis since the criterion (6.35) still excludes some oscillatory approaches to steady state. This means that, if the marginal stability problem is solved for criterion (6.35), a critical Rayleigh number R_c^ζ will be found which is larger than the lesser of R_c^s and R_c^v, but smaller than R_c^{NM}. Note that for $\varepsilon^2 = 0$,

$$R_c^\zeta = R_c^v$$

and for $\tau^2 = 0$,

$$R_c^\zeta = R_c^s \quad . \tag{6.36}$$

Therefore, in order to get the best (largest) value of R_c^ζ for which the system is still stable (not just monotonically stable), R_c^ζ should be maximized with respect to ε^2 and τ^2. However, it is noted that again, in a case where the return to steady state is monotonic, all three criteria are equivalent and

$$R_c^{NM} = R_c^\zeta = \min(R_c^s, R_c^v) \quad . \tag{6.37}$$

In the Bénard problem, it can be proven that there are no oscillatory approaches to the steady state for positive Rayleigh numbers (σ is real). Therefore, in the criterion (6.35), ε^2 and τ^2 may be chosen in any manner, and the result will be the same, namely R_c^{NM}. This is what Glansdorff and Prigogine [1] do. They choose

$$\varepsilon^2 = (T^\dagger)^{-1}\, T^2$$

and

$$\tau^2 = T$$

where

$$T^\dagger = \frac{c_v\, \beta}{\alpha g} \quad . \qquad (6.38)$$

It should be noted that this choice is merely one of convenience and will not affect the resulting value of R_c^ζ because of the monotonic nature of the decays. However, it is true that, in practice, R_c^ζ is often found by a variational method, using some trial function ψ. The value of R_c^ζ for a particular trial function will depend upon ε^2 and τ^2. However, for a completely arbitrary trial function (satisfying boundary conditions, etc.), the extremum of $R_c^\zeta(\psi, \varepsilon^2, \tau^2)$ must be independent of τ^2 and ε^2. The issue is slightly confused by the fact that the actual variational procedure is over the non-trivial, unstable trial functions yielding R_c^ζ as a minimum. That is, R_c^ζ is found as the smallest value of R for which an undamped solution to the perturbation equation can exist.

However, the arguments above are easily modified. For example, in finding R_c^ζ, the minimizing procedure is over a larger function space than in finding R_c^{NM} since some functions which are stable in the normal mode approach (hence excluded from the trial function space) are still unstable in the approach of Glansdorff and Prigogine. Therefore $R_c^{\zeta} \leq R_c^{NM}$ as before.

The above comments on the relative values of R_c^s, R_c^v, R_c^{ζ}, and R_c^{NM} obviously apply only to the true values, and not to the approximate values which may be found by a specific, simple trial function. For example, Glansdorff and Prigogine [1], p. 178, find a value of R_c^{ζ} of approximately 672 for free boundaries which is higher than the exact value of about 657.5, although it is known that the true minimum value of R_c^{ζ} must be less than or equal to R_c^{NM} (actually it is equal to R_c^{NM}).

The fact that all perturbations decay monotonically is the reason that the method of Glansdorff and Prigogine [1] agrees so well with the normal mode method. They are, for the Bénard problem, equivalent. The accuracy of their method with a relatively simple trial function may be largely due to a judicious choice of weighting functions but the "exact" result (for a general trial function) must be independent of this choice.

§6.3 Nature of the Unstable Motion

This theory, although in good agreement with experimental results, is incomplete, even in the range where the approximations are valid, since it does not give a complete description of the instability. The instability is characterized by a typical length λ_c, but nothing is said about the actual shape of the convective cells. For example, there are an infinite number of rectangular shapes all having the same value of λ_c, but with different ratios of side lengths. However, from the symmetry of the system, it might be reasonable to assume that the cells must be regular polygons and that they should tesselate the plane. Thus, the cells should be regular triangles, squares, or hexagons. For each cell pattern, it is then possible to deduce the motion of the fluid in each cell. For example, with regular hexagons, the liquid rises in the center, then flows outward at the top surface and down at the sides.

Actually, it has recently been found possible to predict the structure quite completely in the non-linear regime slightly above the critical Rayleigh number. A number of transitions with different patterns are observed, making the problem rather more complicated than in the simple analysis above. For a review of experimental and theoretical results, see the article by Whitehead [22].

Experimentally, it is observed in a typical case, that immediately after the onset of instability, there is a period of short duration when the cells are moderately regular and are convex polygons of four to seven sides. Eventually, the cells become stable, regular hexagons, properly aligned. In these hexagons, as predicted, the fluid rises in the centers and flows downward at the boundaries between adjacent cells.

CHAPTER VII

STABILITY OF FLUIDS IN MOTION (TURBULENCE)

§7.1 Nature of Turbulent Flows

A final example of a macroscopic system far from equilibrium is supplied by a real (viscous) fluid which is in motion. Even in the simple case where thermal and chemical transports are negligible, the motion of a fluid has many interesting and poorly understood features. The additionally simplified case of an incompressible fluid is almost always considered since the equations of motion are extremely complex without this assumption and it is not a bad approximation to many real liquids under moderate conditions.

A real fluid displays a transition from stable, regular, laminar flow to chaotic, turbulent flow for sufficiently high fluid velocities. It is the nature of this instability and the properties of the turbulent fluid which will be considered. The properties of particular laminar flows are not general, since the precise fluid flow is determined by the individual boundary conditions. Thus, it is not possible to make any general statistical predictions of the fluid properties and each system must be analysed separately. The turbulent flow, however, is essentially random in nature and it would be hoped that the properties of the fluid could be statistically predicted. In order to actually do this, some rather drastic simplifications must be made but still it may be possible to gain some insight into the nature of real turbulent fluids.

Before proceeding, just what is meant by "turbulent flow" will be clarified. Considering a liquid flowing through a pipe for illustration, then it is known that at low velocities the flow is laminar. That is, a fluid element follows a smooth trajectory and an element which was near to it will follow a similar trajectory. Thus, the fluid has "streamlines" of flow which preserve their joint distribution at large distances and long times. This can be differently described by saying that the fluid velocities are strongly correlated at widely separated times. It is also a feature of such laminar flow that the fluid motion can be completely determined from a knowledge of initial and boundary conditions.

If the other parameters of the system are held constant and the fluid velocity is increased, it is found that, at a certain critical velocity, the nature of the motion changes completely. Above the critical velocity, the motion is characterized by strong agitation. The streamlines are chaotically interwoven and rapidly changing. In the Euler coordinate system, this is seen as large and apparently random fluctuations of the fluid velocity field. In the Lagrange system, turbulence results in the rapid mixing of a fluid element with its surroundings. An important feature of turbulent motion is its random nature. It is not possible to determine the fluid motion from a knowledge of the initial and boundary conditions alone. It may only be possible to determine the statistical properties of fluid turbulence. In contrast to laminar flow, fluid velocity correlations die out rapidly for large times, although for short times the requirement of continuity results in large correlations.

The existence of an instability in laminar flow is a direct result of the non-linearity of the equation of motion. The fluid, which is assumed to be incompressible, obeys the Navier-Stokes equation

$$\frac{\partial\vec{v}}{\partial t} + \vec{v} \cdot \vec{\nabla}\vec{v} = -\frac{1}{\rho}\vec{\nabla}p + \nu\nabla^2\vec{v} \tag{7.1}$$

along with the condition for incompressibility

$$\vec{\nabla} \cdot \vec{v} = 0 \quad . \tag{7.2}$$

Eqn. (7.1) is still quite complicated. However, it is possible to understand some features of turbulence without treating it explicitly.

As long ago as 1883, Reynolds realized that a qualitative understanding of the onset of turbulence is possible through simple analysis of the dominant forces. In a fluid, these are the dissipative viscous force $\vec{F}_{vis}$ and the inertial force $\vec{F}_{in}$ which arises from velocity gradients.

The inertial force is of the form

$$|\vec{F}_{in}| \sim |\vec{\nabla}(\rho v^2)| \sim \frac{\rho v^2}{D} \qquad (7.3)$$

where D is a length of the order of the size of the system. The viscous force is of the form

$$|\vec{F}_{vis}| \sim \eta|\nabla^2\vec{v}| \sim \frac{\eta |\vec{v}|}{D^2} \quad . \qquad (7.4)$$

If there exists an initial laminar flow and it is slightly perturbed, the flow will regress to the initial flow if the viscous force is larger than the inertial force. That is, the flow is characterized by a dimensionless Reynolds number Re such that, if Re < 1 the flow is laminar. Here, Re is defined by

$$Re \equiv \frac{|\vec{F}_{in}|}{|\vec{F}_{vis}|} = \frac{\rho\ v^2\ D^{-1}}{\eta\ |\vec{v}|\ D^{-2}} = \frac{|\vec{v}|\ D}{\nu} \quad . \qquad (7.5)$$

In fact, it is found that Re can be larger than one and still allow laminar flow. Instead, there is a critical Reynolds number Re_c such that, for $Re < Re_c$, the viscous force damps the local fluctuations and, for $Re > Re_c$, the local fluctuation is carried away into the fluid as turbulence.

Also, experimentally it is necessary to distinguish between two cases. For the transition from laminar to turbulent flow, it is found that there is no absolute upper limit on Re. Laminar flow has been maintained in circular pipes for Reynolds numbers as high as 50,000, although this flow is highly unstable. For the change from turbulent to laminar flow, it is found that the transition occurs for $Re \approx 1100$ (again in circular pipes) regardless of how the transition is approached.

In this analysis, it is also clear how the random nature of turbulence arises. For $Re > Re_c$, the microscopic fluctuations, which are always present in a fluid, grow rapidly to finite size. The motion thus depends not only upon the macroscopic initial and boundary conditions, but also upon the random microscopic fluctuations.

Another point of view with regard to turbulence is possible by considering velocity correlations. This approach was developed between 1920 and 1941, particularly by Taylor, Richardson, and Kolmogorov. Since velocities in a turbulent fluid are correlated at short times and distances, but not at large times and distances, there is inherent in a turbulent fluid a length scale which is an "average correlation length" of an entirely different scale from the size of the system as a whole. Thus, turbulence can be viewed as the simultaneous existence of macroscopic, cooperative regions characterized by a different order of deviations from the average velocity than the fluid as a whole. These regions are called eddies and have a typical size λ and a typical velocity v_λ. Within the eddies, the velocities are strongly correlated and, outside the eddies, weakly or not at all.

In a real fluid, the eddies are expected to interact with each other and with their surroundings, resulting in the creation, destruction, and transformation of eddies as time progresses.

§7.2 Analytic Description of Turbulence

In this section, for convenience, vectors and tensors will be indicated by indices and summation over repeated indices will be assumed.

The idea of a Reynolds number which determines the nature of the flow can be generalized with the eddy description of turbulence. The fluid velocity can be separated into an average velocity $\overline{v}_i$ and a fluctuating velocity v_i' . That is

$$v_i = \overline{v}_i + v_i' \qquad (7.6)$$

where $\overline{v_i'} = 0$. The energy of the fluctuations is then

$$E' = \frac{1}{2} \rho \, v'^2 \quad . \qquad (7.7)$$

The quantity which determines whether a given flow is turbulent is $\frac{dE'}{dt}$. This can be separated into an inertial and a viscous term as

$$\frac{dE'}{dt} = \frac{dE'}{dt}\bigg|_{in} + \frac{dE'}{dt}\bigg|_{vis} \quad . \qquad (7.8)$$

Then, if $\frac{dE'}{dt} < 0$, there is laminar flow, and if $\frac{dE'}{dt} > 0$, there is turbulent flow. Consider now a localized velocity fluctuation or eddy of size λ and with fluctuating velocity $v'_{\lambda i}$. The kinetic energy associated with the eddy is

$$E' = \frac{1}{2} \rho \; {v'_{\lambda}}^2 \qquad (7.9)$$

and a typical time for the eddy to appear is

$$T_{\lambda} = \frac{\lambda}{v'_{\lambda}} \quad . \qquad (7.10)$$

The energy transformed per unit time to the eddy from the mean motion is

$$\left.\frac{dE'}{dt}\right|_{in} = \frac{\rho \; {v'_{\lambda}}^2}{2T_{\lambda}} \sim \frac{\rho \; {v'_{\lambda}}^3}{\lambda} \quad . \qquad (7.11)$$

On the other hand, the rate of viscous energy dissipation is

$$\left.\frac{dE'}{dt}\right|_{vis} = \frac{\eta}{2}\left(\frac{\partial v'_i}{\partial x_k} + \frac{\partial v'_k}{\partial x_i}\right)^2 \sim \frac{\nu\rho}{2}\left(\frac{v'_{\lambda}}{\lambda}\right)^2 \qquad (7.12)$$

if the velocity gradients in an eddy are assumed to be of order $(\frac{v'_{\lambda}}{\lambda})$. Thus, the condition for an eddy to exist is

$$\frac{\lambda v'_{\lambda}}{\nu} \equiv Re_{\lambda} > 1 \quad . \qquad (7.13)$$

The quantity Re_{λ} is called the internal or spectral Reynolds number. Therefore, there is a minimum size of eddy that can exist, the size of which is denoted by λ_0. Since the eddies are limited by the size of the system, the largest eddy has $\lambda \sim D$. The eddies with the largest Reynolds numbers will be most favourably created and will appear first as turbulence begins. When the fluid velocity is slightly larger than the critical velocity, the turbulence will be dominated by eddies of size D. Smaller eddies may be created, but will soon either die out or grow into larger ones. If the mean fluid velocity becomes larger, the velocities in the large eddies are large enough to allow the creation of smaller secondary eddies. This results in a spectrum of eddies which might typically look like Fig. 7.1.

In the largest eddies, the effect of viscosity is unimportant and energy is lost only through transfer to smaller eddies. This is balanced by the flux of energy into these eddies from the mean motion.

As λ gets smaller, the effect of viscosity becomes more important until, at some size λ_0, the energy is dissipated by frictional forces as quickly as it is received from the larger eddies. The size λ_0 provides an internal length scale which, in atmospheric turbulence, for example, might be about one centimeter, while D might be around a kilometer.

The main problem of turbulence is then to determine the eddy spectrum from the equations of motion. Some of the main approaches to this problem will now be presented.

First, the velocity correlation tensor B_{ij} is defined as

$$B_{ij}(\vec{x},t;\vec{r},\tau) \equiv \overline{v_i'(\vec{x},t)v_j'(\vec{x}+\vec{r},t+\tau)} \qquad (7.14)$$

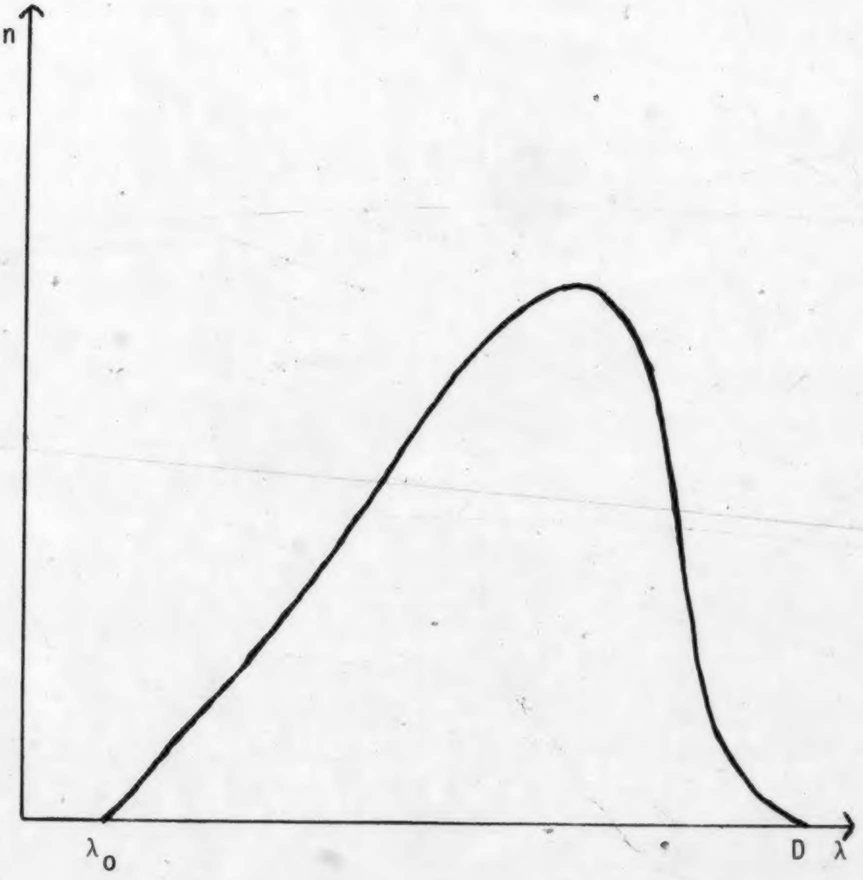


Fig. 7.1 Typical Eddy Spectrum

λ = size of eddies

n = number of eddies

where the bar denotes an average. This then is the average correlation between the fluctuating velocities at two different points in space-time separated by $\vec{r}$ and τ. This tensor is also known as the turbulent energy tensor. Its importance is seen in considering the equation of motion for $\overline{v}_i$. Using eqn. (7.1) and averaging (with $v_i = \overline{v}_i + v_i'$, $p = \overline{p} + p'$, and $\overline{v_i'} = \overline{p'} = 0$) yields

$$\frac{\partial \overline{v}_i}{\partial t} + \frac{\partial}{\partial x_k} \left(\overline{v}_k \overline{v}_i + \overline{v_k' v_i'}\right) = - \frac{1}{\rho} \frac{\partial \overline{p}}{\partial x_i} + \nu \frac{\partial^2 \overline{v}_i}{\partial x_k \partial x_k} \qquad (7.15)$$

and

$$\frac{\partial \overline{v}_k}{\partial x_k} = 0 \quad . \qquad (7.16)$$

These are the Reynolds equations for the mean motion. They consist of four equations in ten unknowns and hence equations are needed for the terms $\overline{v_k' v_i'}$.

If the simplified case where $\overline{v}_i = 0$ is considered (as could be realized experimentally by stirring a fluid violently and then watching the turbulence decay), then the equations for the velocity correlations may be written as follows. First, changing notation slightly, let

$$v_i = v_i(\vec{x},t) \ ; \ v_i' = v_i(\vec{x}',t)$$

and

$$p = p(\vec{x},t) \ ; \ p' = p(\vec{x}',t) \quad . \qquad (7.17)$$

Then, multiplying eqn. (7.1) by v_j' gives

$$v_j' \frac{\partial v_i}{\partial t} + v_j' \frac{\partial (v_i v_k)}{\partial x_k} = - \frac{v_j'}{\rho} \frac{\partial p}{\partial x_i} + \nu\, v_j' \frac{\partial^2 v_i}{\partial x_k \partial x_k} \qquad (7.18)$$

and, similarly,

$$v_i \frac{\partial v'_j}{\partial t} + v_i \frac{\partial (v'_j v'_k)}{\partial x'_k} = - \frac{v_i}{\rho} \frac{\partial p'}{\partial x'_j} + \nu v_i \frac{\partial^2 v'_j}{\partial x'_k \partial x'_k} \quad . \tag{7.19}$$

Adding yields

$$\frac{\partial (v_i v'_j)}{\partial t} + \frac{\partial}{\partial x_k}(v_i v_k v'_j) + \frac{\partial}{\partial x'_k}(v_i v'_k v'_j)$$

$$= - \frac{1}{\rho}\left(\frac{\partial (p v'_j)}{\partial x_i} + \frac{\partial (p' v_i)}{\partial x'_j}\right) + \nu\left(\frac{\partial^2}{\partial x_k \partial x_k} + \frac{\partial^2}{\partial x'_k \partial x'_k}\right) v_i v'_j \; . \tag{7.20}$$

Now considering the case of homogeneous turbulence, the correlations can only depend upon the difference $\vec{r} \equiv \vec{x}' - \vec{x}$ and not upon $\vec{x}$ and $\vec{x}'$ separately. Then, $\frac{\partial}{\partial r_i} = \frac{\partial}{\partial x'_i} = - \frac{\partial}{\partial x_i}$. Also, in the case of isotropic turbulence, it can be shown that $\overline{p\ v'_i} = \overline{p'\ v_j} = 0$. If eqn. (7.20) is averaged it yields

$$\left(\frac{\partial}{\partial t} - 2\nu \nabla_r^2\right) B_{ij}(\vec{r},t) = S_{ij}(\vec{r},t) \tag{7.21}$$

where

$$\nabla_r^2 \equiv \frac{\partial^2}{\partial r_k \partial r_k} \quad ,$$

$$B_{ij}(\vec{r},t) \equiv B_{ij}(\vec{x},t\ ;\ \vec{r},o) \quad ,$$

and

$$S_{ij}(\vec{r},t) \equiv \frac{\partial}{\partial r_k}\left(\overline{v_i v_k v'_j} - \overline{v_i v'_k v'_j}\right) \quad . \tag{7.22}$$

If also $B_{kj,i}$ is defined by

$$B_{kj,i} \equiv \overline{v'_i v'_k v_j} = - \overline{v_i v_k v'_j} \tag{7.23}$$

then

$$\left(\frac{\partial}{\partial t} - 2\nu \nabla_r^2\right) B_{ij} = \frac{\partial}{\partial r_k}\left(B_{ik,j} + B_{jk,i}\right) \quad . \tag{7.24}$$

Taking the trace of eqn. (7.24) gives

$$\left(\frac{\partial}{\partial t} - 2\nu \nabla_r^2\right)B_{ii} = 2\,\frac{\partial}{\partial r_k}\,(B_{ik,i}) \quad . \qquad (7.25)$$

Since the fluid is homogeneous and isotropic, the tensor B_{ij} can be written in terms of two scalar functions of $r(\equiv |\vec{r}|)$.

$$B_{ij}(\vec{r},t) = \frac{1}{r^2}\,[B_{\ell\ell}(r,t) - B_{nn}(r,t)]r_i\, r_j + B_{nn}(r,t)\delta_{ij} \qquad (7.26)$$

where $B_{\ell\ell}$ and B_{nn} are scalar functions corresponding to the projections of $B_{ij}(\vec{r},t)$ along the vector $\vec{r}$ and normal to $\vec{r}$ respectively (no summations over ℓ or n are implied).
For a proof of this, see Panchev [23], pp. 89-93.

Also, for incompressible fluids,

$$\frac{\partial B_{ij}}{\partial r_i} = \frac{\partial\overline{(v_i(\vec{x},t)\;v_j(\vec{x}',t))}}{\partial x_i'} = 0 \qquad (7.27)$$

so

$$\frac{\partial B_{ij}}{\partial r_i} = 0 = \Big[\frac{\partial B_{nn}}{\partial r} + \frac{4}{r}\,(B_{\ell\ell} - B_{nn}) + r^2\,\frac{\partial}{\partial r}\,\Big(\frac{B_{\ell\ell} - B_{nn}}{r^2}\Big)\Big]\,\frac{r_j}{r} \quad . \qquad (7.28)$$

Thus,

$$B_{nn}(r,t) = \frac{1}{2r}\,\frac{\partial}{\partial r}\,(r^2 B_{\ell\ell}(r,t)) \quad . \qquad (7.29)$$

Therefore, $B_{ij}(\vec{r},t)$ can be expressed in terms of the single scalar function $B_{\ell\ell}$ as

$$B_{ij}(\vec{r},t) = -\,\frac{1}{2r}\,\frac{\partial B_{\ell\ell}}{\partial r}\,r_i\, r_j + \Big(B_{\ell\ell} + \frac{r}{2}\,\frac{\partial B_{\ell\ell}}{\partial r}\Big)\delta_{ij} \quad . \qquad (7.30)$$

The function $B_{\ell\ell}(r,t)$ allows one to define a characteristic longitudinal correlation length L_1 as

$$L_1 = \frac{1}{B_{\ell\ell}(o,t)} \int_0^\infty B_{\ell\ell}(r,t)dr \quad . \qquad (7.31)$$

Similarly, it can be shown that the third order correlation $B_{jk,i}$ can be expressed in terms of a single scalar function $B_{\ell\ell\ell}(r,t)$, the longitudinal correlation function of third order.

$$B_{jk,i}(\vec{r},t) = \frac{1}{2r^3}(B_{\ell\ell\ell}-r\frac{\partial B_{\ell\ell\ell}}{\partial r})r_i r_j r_k + \frac{1}{2r}(B_{\ell\ell\ell}+\frac{r}{2}\frac{\partial B_{\ell\ell\ell}}{\partial r}) \times$$

$$(r_j\delta_{ik} + r_k\delta_{ij}) - B_{\ell\ell\ell}\frac{r_i}{2r}\delta_{jk} \quad . \qquad (7.32)$$

Therefore,

$$B_{ik,i}(\vec{r},t) = \frac{1}{2}(\frac{\partial}{\partial r} + \frac{4}{r})B_{\ell\ell\ell}(r,t)r_k \quad . \qquad (7.33)$$

Taking the trace of eqn. (7.30) gives

$$B_{ii}(\vec{r},t) = (r\frac{\partial}{\partial r} + 3)B_{\ell\ell}(r,t) \quad . \qquad (7.34)$$

From the definition of S_{ij},

$$S_{ii}(\vec{r},t) = 2\frac{\partial}{\partial r_k}(B_{ik,i})$$

$$= (r\frac{\partial}{\partial r} + 3)(\frac{\partial}{\partial r} + \frac{4}{r})B_{\ell\ell\ell}(r,t) \qquad (7.35)$$

and it can easily be shown that

$$\nabla_r^2 B_{ii}(\vec{r},t) = (r\frac{\partial}{\partial r} + 3)[\frac{1}{r^4}\frac{\partial}{\partial r}(r^4\frac{\partial B_{\ell\ell}}{\partial r})] \quad . \qquad (7.36)$$

Inserting these expressions in (7.30) gives

$$(r \frac{\partial}{\partial r} + 3)[(\frac{\partial}{\partial t} - \frac{2\nu}{r^4} \frac{\partial}{\partial r} (r^4 \frac{\partial}{\partial r}))B_{\ell\ell}(r,t) -$$

$$(\frac{\partial}{\partial r} + \frac{4}{r})B_{\ell\ell\ell}(r,t)] = 0 \quad . \qquad (7.37)$$

If the expression in the brackets is to be non-singular at r=0, then

$$[\frac{\partial}{\partial t} - \frac{2\nu}{r^4} \frac{\partial}{\partial r}(r^4 \frac{\partial}{\partial r})]B_{\ell\ell}(r,t) = (\frac{\partial}{\partial r} + \frac{4}{r})B_{\ell\ell\ell}(r,t) \quad . \qquad (7.38)$$

This is known as the dynamic equation of Karmán-Howarth. It is an exact equation governing the decay of turbulent correlations in an isotropic, homogeneous, and incompressible fluid. In order to solve for $B_{\ell\ell}(r,t)$ and thus for $B_{ij}(\vec{r},t)$, $B_{\ell\ell\ell}(r,t)$ must be known. If an equation for $B_{\ell\ell\ell}(r,t)$ is derived, it will involve fourth order correlations and so on. Thus, in order to get a solution for the second order correlations, it is necessary to make some assumption about the higher order correlations.

For example, in the final stage of decay, it might be assumed that all triple correlations have disappeared ($B_{\ell\ell\ell}=0$) so eqn. (7.38) becomes

$$\frac{\partial B_{\ell\ell}}{\partial t} = 2\nu \, (\frac{\partial^2}{\partial r^2} + \frac{4}{r} \frac{\partial}{\partial r})B_{\ell\ell} \quad . \qquad (7.39)$$

This equation has a solution of the form

$$B_{\ell\ell}(r,t) \sim \frac{e^{-r^2/8\nu t}}{(\nu t)^{5/2}} \qquad (7.40)$$

near r=0 (i.e. for small eddies). With r=0, the law for decay of turbulent kinetic energy is deduced as

$$\overline{v^2(t)} = B_{ii}(o,t) \sim t^{-5/2} \qquad (7.41)$$

in the final stages of decay. It is also possible to find the time dependence of the correlation length at this stage.

$$L_1(t) = \int_0^\infty e^{-r^2/8\nu t}\, dr \sim t^{1/2} \quad . \qquad (7.42)$$

Thus, the Reynolds number has the dependence

$$Re = \frac{(\overline{v^2(t)})^{1/2}}{\nu} L_1 \sim t^{-3/4} \qquad (7.43)$$

in the final stage of decay.

These power laws have been verified to within the limits of experimental accuracy (cf. Panchev [23]). This gives at least some support to the concepts of homogeneity and isotropy in real fluids, as well as partially verifying the applicability of the Navier-Stokes equation to turbulent fluids.

Rather than dealing directly with the velocity correlations, it is often easier to apply physical intuition in dealing with the Fourier transforms of the correlation functions. This is the basis of the spectral approach to turbulence.

First, define $\Phi_{ij}(\vec{k},t)$ as the Fourier transform of $B_{ij}(\vec{r},t)$:

$$\Phi_{ij}(\vec{k},t) = \frac{1}{(2\pi)^3} \int e^{-i\vec{k}\cdot\vec{r}} B_{ij}(\vec{r},t) d^3\vec{r}. \quad (7.44)$$

The kinetic energy tensor $B_{ij}(o,t)$ can be written as

$$B_{ij}(o,t) = \int \Phi_{ij}(\vec{k},t) d^3\vec{k} \quad . \quad (7.45)$$

Then, from the tensor function $\Phi_{ij}(\vec{k},t)$, a scalar function $\phi(k,t)$ can be defined as the kinetic energy in a spherical shell between k and k + dk .

$$\phi(k,t)dk = \frac{1}{2} \Phi_{ii}(\vec{k},t) \times 4\pi k^2 dk$$

where $k = |\vec{k}|$ and homogeneity and isotropy have been assumed. Then,

$$\frac{1}{2} \overline{v^2(t)} = \int_o^\infty \phi(k,t)dk \quad . \quad (7.46)$$

Also, eqn. (7.21) is transformed to

$$\frac{\partial}{\partial t} \Phi_{ij}(\vec{k},t) + 2\nu k^2 \Phi_{ij}(\vec{k},t) = \Gamma_{ij}(\vec{k},t) \quad (7.47)$$

where
$$\Gamma_{ij} = ik_m(\Phi_{im,j} + \Phi_{jm,i}) \quad (7.48)$$

and $\Phi_{im,j}$ is the Fourier transform of $B_{im,j}$.

By taking the trace of eqn. (7.47) the equation for $\phi(k,t)$ is found as

$$\frac{\partial\phi}{\partial t}(k,t) + 2\nu k^2 \phi(k,t) = W(k,t) \quad (7.49)$$

where
$$W(k,t) = 2\pi k^2 \Gamma_{ii}(\vec{k},t) \quad . \quad (7.50)$$

If eqn. (7.49) is written as

$$\frac{\partial \phi}{\partial t} = W - 2\nu k^2 \phi \tag{7.51}$$

then W may be identified as the rate of energy transfer from the large to the small eddies, and the term $-2\nu k^2\phi$ as the rate of viscous dissipation of energy in eddies of wave number k. It is easily seen that the transfer of energy between different sizes of eddies depends only upon the third order correlations.

It is possible to show that

$$\int_0^\infty W(k,t)\, dk = 0 \tag{7.52}$$

which means that the interaction between eddies does not affect the total viscous energy dissipation.

In the final stage of decay, $W \approx 0$, since the third order correlations disappear. Then,

$$\phi(k,t) = \phi(k,o)e^{-2\nu k^2 t} \tag{7.53}$$

which again shows that the eddies with the largest k (smallest size) are most quickly damped by viscosity.

The basic problem then could be stated as finding the functional dependence of the energy transfer W upon the spectrum ϕ. Once this is known, it is possible to solve for ϕ and essentially everything about the turbulence is known.

The general nature of the spectrum and the energy transfer and loss is summarized in Fig. 7.2.

In the range $k < K_0$, the eddies are produced from the mean motion of the fluid. For $K_0 < k < k_0$, the main mechanism is transfer of energy from large to small eddies. This is referred to as the inertial range. For $k > k_0$, viscous forces dominate.

The functional form of $\phi(k,t)$ can be determined for $K_0 < k < k_0$ by dimensional analysis. In this region, viscous dissipation and eddy creation from the mean motion are negligible. Thus, the statistical nature of the turbulence is determined by a constant parameter ε which determines the energy transfer from large to small eddies. Hence, ϕ is a function only of ε and k . If a power law of the form

$$\phi(k,t) = \alpha\ \varepsilon^p\ k^q \tag{7.54}$$

is assumed, then, since the dimensions of ϕ are (length)^3 (time)$^{-2}$, it is found that $p = 2/3$ and $q = 5/3$. This means that

$$\phi(k,t) = \alpha\ \varepsilon^{2/3}\ k^{-5/3} \quad . \tag{7.55}$$

In order to gain some insight into the dependence of W upon ϕ, integrate the equation of motion (7.51) from zero to k, giving

$$\frac{\partial}{\partial t}\int_0^k \phi(p,t)dp = \int_0^k W(p,t)dp - 2\nu \int_0^k \phi(p,t)p^2\ dp. \tag{7.56}$$

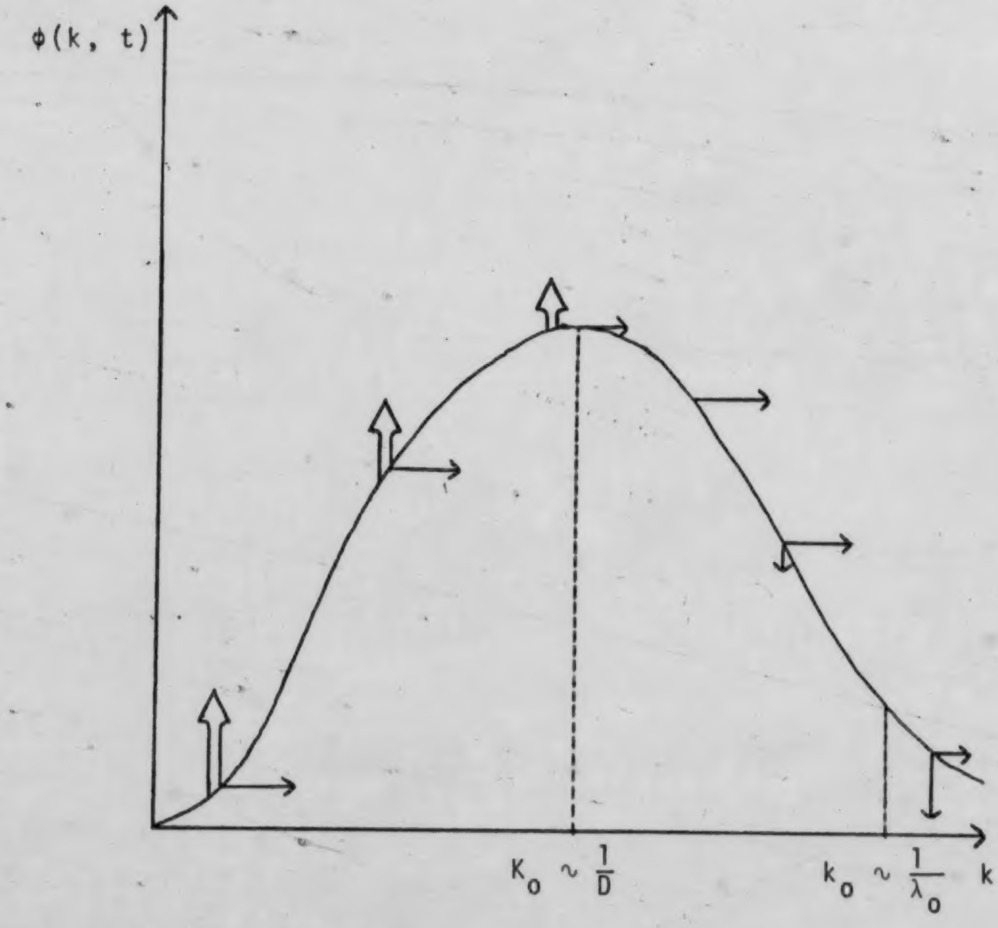


Fig. 7.2 Turbulent spectrum and energy transfer

⇧ : eddy production from mean motion

→ : energy transfer from large to small eddies

↓ : viscous dissipation of kinetic energy

Since most of the kinetic energy is in the small eddies for which $p \sim K_0$,

$$\int_0^k \phi(p,t)dp \approx \int_0^\infty \phi(p,t)dp \tag{7.57}$$

for $k >> K_0$. Then,

$$\frac{\partial}{\partial t}\int_0^k \phi(p,t)dp \approx \frac{\partial}{\partial t}\int_0^\infty \phi(p,t)dp = -\varepsilon\ . \tag{7.58}$$

Thus,

$$\varepsilon = F(k) + 2\nu \int_0^k \phi(p,t)p^2\, dp \tag{7.59}$$

where $$F(k) = -\int_0^k W(p,t)dp = \int_k^\infty W(p,t)dp \tag{7.60}$$

Here, $F(k)$ is an energy flux function which describes the transfer of energy from large to small eddies.

In order to extract further information about the turbulent motion, it is necessary to assume a specific form for $F(k)$. This means that an assumption about the dominant mechanism of energy transfer must be made. Many such assumptions have been made (cf. Panchev [23], pp. 198-219). Here, only one will be described - the ansatz of Heisenberg.

For viscous dissipation in the smallest eddies, it can be shown that the rate of dissipation, ε, is given by

$$\varepsilon = \nu\, \overline{|\vec{\nabla} \times \vec{v}|^2} \tag{7.61}$$

and

$$\overline{|\vec{\nabla} \times \vec{v}|^2} = 2 \int_0^\infty \phi(p)\, p^2\, dp$$

$$\approx 2 \int_0^k \phi(p)\, p^2\, dp \qquad (7.62)$$

for large k. Heisenberg assumed that the process of transfer of energy from large to small eddies is similar to the process of viscous dissipation in small eddies. That is, he assumed

$$F(k) = \nu_T(k) \int_0^k 2\phi(p)\, p^2\, dp \qquad (7.63)$$

where $\nu_T(k)$ is a "viscosity" coefficient describing the damping of large eddies due to the motion of small eddies.

Thus, the spectral equation becomes

$$\varepsilon = [\nu + \nu_T(k)] \int_0^k 2\phi(p)\, p^2\, dp \quad . \qquad (7.64)$$

In order to find a plausible form of $\nu_T(k)$, assume that it depends upon ϕ and k:

$$\nu_T(k) = \nu_T[\phi, k] \qquad (7.65)$$

where ν_T has units of $(\text{length})^2 \div (\text{time})$. Since the damping of large eddies should depend on the presence of any smaller eddies, it is expected that $\nu_T(k)$ is a function of all $p > k$. From dimensional arguments,

$$\nu_T(k) = \gamma_1 \int_k^\infty \phi^{1/2}(p)\, p^{-3/2}\, dp \qquad (7.66)$$

which is Heisenberg's original expression for ν_T. For $k \to \infty$,

$$2 \int_0^k \phi(p)\, p^2\, dp \approx \frac{\varepsilon}{\nu} \tag{7.67}$$

so, as $k \to \infty$,

$$F(k) \to \nu_T(k) \cdot \frac{\varepsilon}{\nu} \quad . \tag{7.68}$$

Therefore,

$$k^2\, \phi(k) \to \frac{\varepsilon}{2\nu^2} \frac{d\, \nu_T(k)}{dk} \quad . \tag{7.69}$$

Substituting expression (7.66) for ν_T gives

$$\phi(k) \to \frac{1}{4} \frac{\varepsilon^2 \gamma_1^2 k^{-7}}{\nu^4} \tag{7.70}$$

as $k \to \infty$.

Unfortunately, the Heisenberg ansatz suffers from serious drawbacks. First, although the idea of small eddies acting as an effective viscosity for large ones seems plausible, there is no reason to suppose that this is a good description of the interaction between eddies of nearly the same size. Also, the dependence $\nu_T = \nu_T[\phi, k]$ is not consistent with the physical requirement that the small eddies be statistically independent of the large ones.

The Heisenberg expression for $W(k)$ agrees only qualitatively with experimental results, and there are some features which are not experimentally observed (cf. Uberoi

[24]). Although part of the difficulty may be in obtaining the necessary isotropy and homogeneity in a real fluid, it would seem that energy transfer is a complex phenomena, unlikely to be accurately described by such a simple ansatz as that of Heisenberg.

PART II

CHAPTER VIII

INTRODUCTION AND FOUNDATIONS

§8.1 Historical Introduction

Although macroscopic thermodynamics was a well established branch of physics by the middle of the nineteenth century, there were many aspects of thermodynamics which were essentially phenomenological. For example, thermodynamics was incapable of quantitatively describing the behavior of a specific material. For example, the equation of state of a given material was not a theoretical deduction from thermodynamics but rather was an experimental addition to the theory. This deficiency was, of course, remedied with the establishment of equilibrium statistical mechanics. A similar situation existed in non-equilibrium thermodynamics where the coefficients relating the thermodynamic forces and fluxes had to be taken from outside the theory.

In order to develop a theory which could yield both equations of state and transport coefficients it was necessary to know what determined the macroscopic properties of different materials. The essential development which permitted such a general theory was the acceptance of the atomic theory of matter. This theory, which was based largely upon the study of simple chemical reactions in gases, resulted in the point of view that the macroscopic properties of a material depend only upon the properties of the identical atoms that comprise the material.

Although the atomic theory itself was controversial throughout the second half of the nineteenth century, it had a number of important applications in the kinetic theory of gases during this period.

This microscopic theory resulted in the derivation of equations of state and transport properties of dilute gases. As well as these practical results, such theories of gases had more profound effects. The kinetic theories of Maxwell, Boltzmann, and others gave a microscopic, mechanical interpretation of such established thermodynamic concepts as temperature, pressure, and chemical potential. In so doing, they pointed the way to generalizations and further applications. Also, from the beginning, kinetic theories introduced the concept of probability. At first, probability was introduced as an assumption that all the molecules moved independently or in only one direction or something similarly simple, and was seen as an undesirable aspect of the theory necessary only because of the complexity of the dynamical equations. Later, the probability aspect became both more sophisticated and more firmly entrenched as it was realized that the extremely large number of molecules in a macroscopic system made the dynamical solution impossible and the statistical solution more precise.

Despite the successes of the kinetic theory, there was considerable resistance to the statistical theory throughout the latter part of the nineteenth century. This was in large part due to a feeling that probability concepts were foreign to a previously deterministic physics and must eventually be supplanted by an "exact" and "correct" theory.

However, as the evidence for the atomic nature of matter mounted, it became more obvious that statistical concepts were not just temporarily necessary and were not limited in applicability to dilute gases. One of the most compelling arguments for both the atomic theory of matter and for the applicability of statistics to systems other than dilute gases was the remarkable success of the statistical theory of Brownian motion. Brownian motion, which provides almost direct evidence of the existence of molecules, was analysed using statistical methods by Einstein and by von Smoluchowski in the first few years of the twentieth century. In addition to this evidence of the usefulness of statistical methods, the acceptance of these concepts was probably accelerated by the publication by Gibbs, in 1902, of an elegant and precise formulation of statistical physics. In his treatise, Gibbs introduced the concept of ensembles of identical systems and, in so doing, separated the concept of probability from the deterministic dynamics of a system.

The combination of the above factors resulted, by the end of the first decade of this century, in the general acceptance of statistical physics as a major branch of theoretical physics. The many successes of the theory have since justified this acceptance. In addition, statistical physics easily made the transition from classical to quantum mechanics.

With these historical comments, this chapter will attempt to examine some of the fundamental concepts of statistical physics. After having introduced the main concepts, further chapters will give more detailed applications of the statistical method to both classical and quantum non-equilibrium systems. The emphasis, however, will be on general formalisms and examples illustrating general principles rather than upon detailed calculations for special systems. Also, no attempt will be made to exhaustively cover the huge range of non-equilibrium situations.

§8.2 Classical Foundations

Classically, there are two reasons for introducing probability concepts into the physics of systems with a large number of degrees of freedom. The first is the practical consideration that it is impossible to actually solve the exact dynamical equations for complex systems. Thus, it is hoped that, by ignoring the details of the actual motion, simpler equations may be obtained for the evolution of averaged quantities. However, this alone is not a satisfactory justification for introducing probability concepts since there is no intuitive reason why it should always be possible to describe a nearly arbitrary motion in terms of a few statistical parameters. On the other hand, even if it were possible to exactly solve Hamilton's equations, thus eliminating the above reason for introducing probabilities, the resultant evolutions would still depend on the initial conditions. Then, since it is impossible, from a practical point of view, to measure simultaneously the initial position and momentum of each particle, the initial condition must be described as a statistical distribution, rather than a single deterministic quantity. This point of view is even more reasonable when non-isolated systems are considered since the thermal interaction with the surroundings will act to randomize the motion of the system. With this point of view, there is no conceptual difficulty in introducing probabilistic ideas since they arise from an initial lack of knowledge about the system and, once introduced, the statistical distribution evolves in a deterministic way governed by Hamilton's equations.

In quantum mechanics, this uncertainty in initial conditions enters due to the practical impossibility of determining a complete set of eigenvalues.

Of course, no matter how the probabilities are introduced, there is still the practical difficulty of determining or guessing the distribution itself. This often takes the form of an assumption of equal a priori probabilities or something similarly simple.

It is the introduction of probabilistic initial conditions that corresponds to the representative ensembles of Gibbs. If a large number of identical replicas of the system are considered (e.g. identical with respect to such measureable macroscopic parameters as energy or momentum), then this collection of systems is said to form a representative ensemble. As the number of copies of the system tends to infinity, the members of the ensemble (considered at $t = 0$) will distribute themselves according to the probability distribition $\rho(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N, t=0)$ over the microscopic initial conditions on $(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N)$. Here the coordinate $\vec{x}_i$ represents the three position and three momentum coordinates of the i^{th} particle and $\rho(\vec{x}_1,\ldots,\vec{x}_N,t=0)\ d^6\vec{x}_1\ldots d^6\vec{x}_N$ is the probability of finding the system initially in the volume $d^6\vec{x}_1\ldots d^6\vec{x}_N$ of the 6N dimensional Γ-space of the system. To actually follow the time evolution of ρ corresponds to exactly solving Hamilton's equations for all possible initial conditions.

It is realized that, in practice, a particular member of the ensemble is chosen and its time development followed. The process of measuring some dynamical parameter corresponds to taking an average of that parameter over a finite time τ. It is obviously desirable to determine the conditions under which this finite time average is equal to the average over the ensemble. Most of the effort in this area has been directed toward equilibrium situations and this special case is reasonably well understood. Proofs of the equality of time and ensemble averages involve what is known as the ergodic theorem. Since the main interest in this thesis in in non-equilibrium situations, only the basic results of ergodic theory will be mentioned (cf. Isihara [25]).

Since, in equilibrium, the relevant macroscopic parameters are time independent, the averages over a finite time τ may be replaced by averages over an infinite time. It can be proven that, for almost any initial conditions, the limit ($\tau \to \infty$) of the time averaged parameters exists, provided that the phase space is metrically indecomposable and also that, under this condition, the time average is equal to the average over the initial ensemble. Metric indecomposability means that the phase space cannot be split into two invariant parts of finite measure such that a motion starting in one part remains in that part forever. Under the assumption of metric indecomposability, any trajectory, whatever its initial conditions, will pass arbitrarily close to any point on the surface of constant energy (called the ergodic surface) given sufficient time. This means that the average of a function over the entire trajectory (infinite time average) is the same as the average over an ensemble in which the member systems are distributed uniformly over the ergodic surface.

Thus, the problem of justifying representative ensembles in equilibrium statistical mechanics involves the assumptions that actual measurements correspond to infinite time averages and that the phase space for the system is metrically indecomposable. The first is plausible for equilibrium systems, provided the time scale for measurements is much larger than that for fluctuations. The second is a rather abstract mathematical concept which requires investigation for particular systems. Suffice to say that the phase space of certain simple physical systems, such as a gas of hard spheres in a box, has been shown to be metrically indecomposable.

Since proofs of the equivalence of time and ensemble averages depend critically upon the averages being over infinite times, the theory does not apply to non-equilibrium situations. It is precisely the information on time-development which is of interest in most non-equilibrium situations and this information is lost in averaging over all times. It may be that, in systems where a microscopic relaxation time exists which is sufficiently small compared to the macroscopic evolution time of the system, an analogue to the equilibrium ergodic theorem exists. This theorem should indicate the conditions under which finite time averages would be approximately equivalent to averages over an ensemble. However, such a theorem would require some use of the dynamics of the particular system in order to determine the relaxation and evolution times. Due to these basic difficulties, the usual approach in non-equilibrium statistical physics is not via an ergodic theorem.

Instead, it is usual to treat the time development of an arbitrary initial distribution and to simply assume that averages over the ensemble at a particular instant are equivalent to physical measurements on a single system at that instant. The justification of this essentially rests on the observation that, for most physical systems, repeating a non-equilibrium experiment starting with the same macroscopic initial conditions gives the same macroscopic evolution even though the system must have vastly different microscopic initial conditions in each repetition of the experiment. Thus, it seems plausible that, in such situations, the ensemble average could be identified with the macroscopic measurement since the macroscopic evolution is not sensitive to the precise initial conditions. Of course, it is expected that there are some situations where this is no longer true and in which this statistical method is not valid, but it is hoped that the common non-equilibrium situations are not of this perverse nature.

In practice, either a plausible initial distribution is assumed and its development followed, or else an attempt is made immediately to simplify the description by considering reduced probability distributions. The former approach is used, for example, in linear response theory where the initial ensemble is assumed to be an equilibrium one and its time evolution under the influence of an external perturbation is considered. In the latter approach, upon which the Boltzmann and other kinetic equations are based, rather drastic assumptions are made regarding the relationships between the reduced distributions and the complete distribution function.

Thus, although in principle the development of any initial probability distribution can be found and then macroscopic averages may be taken over the resulting time-dependent ensemble, in practice this is not often possible. This is due to the lack of any simplifying feature such as that provided in the equilibrium theory by the time independence of the ensembles. The assumptions which must be made in order to obtain concrete results put much of the non-equilibrium theory on a much less secure footing than the equilibrium theory. In many cases, the only test as to the validity of the statistical assumptions made is comparison of predictions with experiment. Despite this, there are many useful and valid results in the classical statistical theory of non-equilibrium systems. Some of these will be presented in the next two chapters.

§8.3 Quantum Mechanical Formulation

In quantum mechanics, the statistical distribution operator or density matrix that is analogous to the classical distribution function $\rho(\vec{x}_1,\ldots\vec{x}_N,t)$ may be introduced in two fundamentally different ways (cf. ter Haar [26]). The first method is due to von Neumann and is essentially the same as the introduction of the classical distribution as discussed above. To reformulate the classical theory in quantum mechanical terms, proceed as follows.

If the state of an isolated quantum mechanical system is known as completely as possible, then it is said to be in a "pure" state. The system may then be completely described by a wave function ψ. The expectation value $\overline{A}$ of an operator A is then given by

$$\overline{A} = (\psi, A\psi) \quad . \tag{8.1}$$

The wave function $\psi(\vec{x}_1,\ldots,\vec{x}_N,t)$ may be expanded in terms of a complete set of time-independent, orthonormal functions $\{\phi_n\}$. That is

$$\psi(\vec{x}_1,\ldots,\vec{x}_N,t) = \sum_n c_n(t)\, \phi_n(\vec{x}_1,\ldots,\vec{x}_N) \tag{8.2}$$

where all the time dependence is contained in the coefficients $c_n(t)$. With this expansion,

$$\overline{A} = \sum_{n,m} c_m^{\star} c_n (\phi_m, A\, \phi_n) \quad . \tag{8.3}$$

The pure state corresponds to an exactly specified initial state in classical mechanics and the only statistical ideas introduced so far have been in the purely quantum mechanical expectation values.

On the other hand, just as in classical mechanics, practical considerations limit the available information about the system. Only a few parameters of the system are known and the values of the other physical observables are unknown. Thus, the natural procedure in quantum mechanics is also to introduce a representative ensemble to represent the uncertainty in initial conditions. If the members of the ensemble are labelled by the index "k" and have wave functions ψ^k, then the ensemble average $<\overline{A}>$ for an ensemble of N system is just

$$\begin{aligned} \langle\overline{A}\rangle &= \frac{1}{N} \sum_{k=1}^{N} (\psi^k, A\,\psi^k) \\ &= \frac{1}{N} \sum_{k=1}^{N} \sum_{m,n} c_m^{k*}(t)\, c_n^k(t)\, (\phi_m, A\,\phi_n) \quad . \qquad (8.4) \end{aligned}$$

If an operator ρ, called the density operator or matrix, is introduced, defined by

$$\rho_{nm} = (\phi_n, \rho\,\phi_m) = \frac{1}{N} \sum_{k=1}^{N} c_m^{k*}(t)\, c_n^k(t) \qquad (8.5)$$

then

$$\begin{aligned} \langle\overline{A}\rangle &= \sum_{m,n} (\phi_m, A\,\phi_n)(\phi_n, \rho\,\phi_m) \\ &= \mathrm{Tr}[\rho A] \quad . \qquad (8.6) \end{aligned}$$

Since the trace operation is invariant under cyclic permutations, the ensemble averages are invariant under unitary transformations of the basis set $\{\phi_n\}$. Also, note that ρ is Hermitian and that $\mathrm{Tr}\rho = 1$.

In the language of quantum mechanics, the general situation considered above is a "mixed" state. A mixed state cannot be described by a single wave function, but has some probability of being in different states. The diagonal entries of the density operator give the normalized probabilities that a given state ϕ_n is realized at time t.

Such a density operator has not introduced any fundamentally new ideas into statistical mechanics. It resulted from considering representative ensembles with the main difference from the classical case being the replacement of exact values of physical parameters by quantum expectation values. As in classical mechanics, it remains to specify an initial value of the density matrix by other considerations. Once this is done, the time development of ρ may, in principle, be determined from the quantum mechanical version of the Liouville equation which governs the classical distribution function. This equation is known as the von Neumann equation.

It is worth noting that an approach via the Liouville equation in classical mechanics or the von Neumann equation in quantum mechanics is valid only for isolated systems, since these equations are deterministic, dynamical equations for isolated systems and only the initial conditions are statistically distributed. Although it might be expected that non-isolated equilibrium systems would behave in a similar manner to isolated systems with the same energy (e.g. equivalence of canonical and microcanonical ensembles), there is no real justification for such an expectation in non-equilibrium systems. This point is important when considering the formal theory of linear and non-linear response in quantum mechanics (cf. Chapter XI).

There is another way to introduce the density matrix which does not involve the use of ensembles as a starting point. This approach, which is sometimes called the quantum mechanical approach to the density matrix as opposed to the statistical approach above, is due to Dirac (1930). It has the advantage that it naturally deals with non-isolated systems as well as isolated ones. Its disadvantage is that it does not provide an equation of motion of the system's density matrix without a detailed knowledge of the system's surroundings.

As an example, consider a many-electron atom. This atom may be assumed to be in a pure state, described by a wave function Ψ. However, even for a two-electron atom, it has not been possible to find an exact analytic form for this wave function. For more complex atoms, even finding numerical approximations to the wave function is very difficult. On the other hand, if only the energy of such an atom is of interest, the approach via the total wave function is wasteful. Since all the terms in the Hamiltonian involve coordinates of either one or two particles, all the averages of interest will be over at most two particles at a time. Thus, if the one and two body reduced density matrices are defined by integrating $\Psi^*\Psi$ over all but one or two of the coordinates, then a knowledge of these reduced density matrices is sufficient to determine the energy of the total atom exactly. It may be said that the rest of the electrons act as a "heat bath" which introduces a statistical distribution of the one and two electron reduced density matrices, even though the total system is in a pure state. Although the reduced density matrices are useful in quantum chemistry in precisely the above form, the idea may be generalized to thermodynamic systems.

In a system with many degrees of freedom, only a few of these are usually of interest. Collectively denoting those of interest by x, and the rest by q, the total system has a single wave function $\Psi(x,q)$. The reduced density matrix ρ is defined by

$$(x,\rho x') \equiv \int \Psi^*(x',q)\, \Psi(x,q)dq \quad . \qquad (8.7)$$

The quantum mechanical average of an operator A, which depends only on the coordinates x, is

$$\begin{aligned} \langle A\rangle &= \iint \Psi^*(x,q)\ A\ \Psi(x,q)dxdq \\ &= \iint (x',\rho x)(x,Ax')dxdx' \\ &= \mathrm{Tr}[\rho A] \quad . \qquad (8.8) \end{aligned}$$

Consider the case of a system S consisting of a subsystem S_1 which is of interest and another S_2 which is not. Then, integrating over the coordinates of S_2 gives the reduced density matrix for S_1. In the special case where the subsystems do not interact, the total wave function is a product of the wave functions of the subsystems, so the reduced density matrix for S_1 corresponds to a pure state of S_1. If, however, the subsystems interact, the subsystem S_1 cannot be described by a single wave function and the reduced density matrix corresponds to a statistical description of S_1. The subsystem S_2 may be made very large and thus may be thought of as a thermal reservoir for S_1. It is then the interaction with this reservoir which gives rise to the statistical distribution of the reduced density matrix even though the total system is in a well defined quantum state.

Although this gives an intuitively satisfying explanation of the introduction of density matrices and representative ensembles in quantum statistics, it does not solve the practical problem of how to find the density matrices. In order to rigorously justify an assumption about the form of the density matrix of a particular system, the complete solution of Schrödinger's equation for the system plus its surroundings must be known. Since this is almost never possible, the above approach is not practical for large systems and an ad hoc assumption is usually made about the initial density matrix.

This concludes the discussion of the general foundations of the statistical method. The limitations of the theory should be kept in mind in succeeding chapters but generally will not be mentioned explicitly.

CHAPTER IX

CLASSICAL STATISTICAL MECHANICS AND KINETIC THEORY

§9.1 Dynamical Equations

Although it is certain that the correct approach to any mechanical or statistical mechanical problem is through application of quantum mechanics, it is still worthwhile to study the general statistical mechanical problem by purely classical methods. One reason is that there is a wide range of physical systems in which quantum effects are negligible. These include dilute systems, high temperature systems, and much of macroscopic hydrodynamics and continuum mechanics. Also, it is possible to derive many formal methods of solution and some exact results which are easily modified in the quantum case. Historically, some of these methods and results are quite new, having been derived well after the establishment of quantum physics. In addition, many general arguments about the nature of irreversibility apply equally well to the classical and quantum cases, although they were formulated well before the establishment of quantum mechanics. This aspect will be discussed in more detail in the next chapter.

In any closed mechanical system, the classical equations of motion are Hamilton's dynamical equations.

However, for statistical mechanical purposes, it is more convenient to cast Hamilton's equations as an equation for the distribution function ρ. This equation, known as the Liouville equation, may be derived as follows. Consider a system of N particles. Its motion may be described as the motion of a single point in a 6N dimensional Γ-space, obeying Hamilton's equations,

$$\dot{r}_i^\alpha = \frac{\partial H_N}{\partial p_i^\alpha}$$

$$\dot{p}_i^\alpha = - \frac{\partial H_N}{\partial r_i^\alpha} \qquad (i=1,2,\ldots,N\ ;\ \alpha=1,2,3) \qquad (9.1)$$

where the dot denotes the time derivative.

If now an ensemble of identical, independent systems, with a density ρ of representative points in Γ-space, is considered, then, since members of the ensemble are neither created nor destroyed, the density ρ obeys the continuity equation (cf. §2.1)

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}_x \cdot (\dot{\vec{x}}\rho) = 0 \qquad (9.2)$$

where ($\vec{\nabla}_x \cdot$) represents the 6N dimensional divergence and $\vec{x}$ represents the 6N coordinates ($\vec{r}_1, \vec{p}_1, \vec{r}_2, \ldots, \vec{p}_N$). Writing this as

$$\frac{\partial \rho}{\partial t} + \dot{\vec{x}} \cdot \vec{\nabla}_x \rho + \rho\, \vec{\nabla}_x \cdot \dot{\vec{x}} = 0 \qquad (9.3)$$

and using Hamilton's equations gives

$$\vec{\nabla}_x \cdot \dot{\vec{x}} = \sum_{i=1}^{N} \left(\frac{\partial \dot{\vec{r}}_i}{\partial \vec{r}_i} + \frac{\partial \dot{\vec{p}}_i}{\partial \vec{p}_i}\right)$$

$$= \sum_{i=1}^{N} \left(\frac{\partial^2 H_N}{\partial \vec{r}_i \partial \vec{p}_i} - \frac{\partial^2 H_N}{\partial \vec{p}_i \partial \vec{r}_i}\right)$$

$$= 0 \quad . \tag{9.4}$$

Therefore,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{N} (\dot{\vec{r}}_i \cdot \vec{\nabla}_{r_i} \rho + \dot{\vec{p}}_i \cdot \vec{\nabla}_{p_i} \rho) = 0 \quad . \tag{9.5}$$

Again using Hamilton's equations gives Liouville's equation

$$\frac{\partial \rho}{\partial t} = \{H_N, \rho\} \tag{9.6}$$

where {A,B} is the Poisson bracket of A and B, defined by

$$\{A,B\} = \sum_{i=1}^{N} [(\vec{\nabla}_{r_i} A)\cdot(\vec{\nabla}_{p_i} B) - (\vec{\nabla}_{p_i} A)\cdot(\vec{\nabla}_{r_i} B)] \quad . \tag{9.7}$$

The Liouville equation, although exact, is of limited value, since it is far too complex to solve in its entirety for realistic physical systems. Instead, the usual course is to make some approximation which involves the loss of some of the dynamical information about the system, yet which allows a reasonable determination of the quantities of interest. A systematic procedure for doing just this is embodied in the BBGKY hierarchy of equations (introduced by Bogolyubov, Born and Green, Kirkwood, and Yvon, independently). Before deriving this hierarchy, reduced distribution functions will be defined. First, consider a function $f^{(N)}(\vec{x}_1,\ldots\vec{x}_N,t)$, proportional to ρ, such that

$$\int_\Gamma f^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,t)d^6\vec{x}_1\ldots d^6\vec{x}_N = V^N \qquad (9.8)$$

where $\vec{x}_i = (\vec{r}_i, \vec{p}_i)$ and the integral extends over all Γ-space. Also, define the reduced distribution function of order ℓ by

$$f^{(\ell)}(\vec{x}_1,\ldots\vec{x}_\ell,t) = V^{(\ell-N)} \int f^{(N)}(\vec{x}_1,\ldots\vec{x}_N,t)d^6\vec{x}_{\ell+1}\ldots d^6\vec{x}_N \qquad (9.9)$$

where again the integrations are over the whole range of $\vec{x}_i$.

Since the particles are identical, ρ and $f^{(N)}$ are symmetric functions of the $\vec{x}_i$ and so the reduced distribution function $f^{(\ell)}$ is also symmetric in the $\vec{x}_i$'s, independent of the particular choice of ℓ particles. Then, the expression

$$V^{-\ell} f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,t)d^6\vec{x}_1\ldots d^6\vec{x}_\ell \qquad (9.10)$$

gives the probability that any ℓ particles are located in the volumes $d^6\vec{x}_1, d^6\vec{x}_2,\ldots d^6\vec{x}_\ell$ about the points $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_\ell$ at time t.

In order to derive equations for the evolution of the $f^{(\ell)}$, integrate Liouville's equation over all but ℓ of the coordinates. This procedure yields an exact formal hierarchy for the reduced distribution functions. Once this hierarchy is found, some reasonable approximation to one of the distribution functions may be made, yielding approximate equations for the lower order distribution functions.

Consider a system of N identical particles in a volume V, influenced by an external potential $\Phi(\vec{r}_i)$, and interacting via an interparticle potential $\phi_{ij}(|\vec{r}_i-\vec{r}_j|)$ (denoted simply by ϕ_{ij}). The Hamiltonian is then

$$H_N = \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} + \Phi(\vec{r}_i)\right) + \sum_{1\leq i<j\leq N} \phi_{ij}(|\vec{r}_i-\vec{r}_j|) \quad . \quad (9.11)$$

Note that, since $f^{(N)}$ is normalized, it must vanish sufficiently rapidly at infinity that

$$\begin{aligned}\int\left\{\left(\frac{p_i^2}{2m} + \Phi(\vec{r}_i)\right), f^{(N)}\right\} d^6\vec{x}_i &= \sum_{i=1}^{N} \int \left(\vec{\nabla}_{r_i}\Phi \cdot \vec{\nabla}_{p_i} f^{(N)} - \frac{\vec{p}_i}{m} \cdot \vec{\nabla}_{r_i} f^{(N)}\right) d^3\vec{r}_i\, d^3\vec{p}_i \\ &= \sum_{i=1}^{N} (\vec{\nabla}_{r_i}\Phi) f^{(N)} \Bigg|_{\vec{r}_i=-\infty}^{\vec{r}_i=\infty} - \sum_{i=1}^{N} \frac{\vec{p}_i}{m} f^{(N)} \Bigg|_{\vec{p}_i=-\infty}^{\vec{p}_i=\infty} \\ &= 0 \qquad (9.12)\end{aligned}$$

and similarly

$$\int \{\phi_{ij}, f^{(N)}\} d^6\vec{x}_i d^6\vec{x}_j = 0 \quad . \qquad (9.13)$$

Thus, integrating Liouville's equation over $(N-\ell)$ coordinates gives

$$\begin{aligned}
\frac{\partial f^{(\ell)}}{\partial t} &= V^{\ell-N} \frac{\partial}{\partial t} \int f^{(N)}(\vec{x}_1, \ldots, \vec{x}_N, t) d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N \\
&= V^{\ell-N} \int \{ \sum_{i=1}^{N} (\frac{p_i^2}{2m} + \Phi(\vec{r}_i)), f^{(N)}\} d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N \\
&\quad + V^{\ell-N} \int \{ \sum_{1\leq i<j\leq N} \phi_{ij}, f^{(N)}\} d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N \\
&= V^{\ell-N} \int \{ \sum_{i=1}^{\ell} (\frac{p_i^2}{2m} + \Phi(\vec{r}_i)), f^{(N)}\} d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N \\
&\quad + V^{\ell-N} \int \{ \sum_{1\leq i<j\leq \ell} \phi_{ij}, f^{(N)}\} d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N \\
&\quad + V^{\ell-N} \int \sum_{i=1}^{\ell} \sum_{j=\ell+1}^{N} (\vec{\nabla}_{r_i} \phi_{ij}) \cdot (\vec{\nabla}_{p_i} f^{(N)}) d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N \\
&= \{[\sum_{i=1}^{\ell} (\frac{p_i^2}{2m} + \Phi(\vec{r}_i)) + \sum_{1\leq i<j\leq \ell} \phi_{ij}], f^{(\ell)}\} \\
&\quad + (N-\ell) V^{\ell-N} \int \sum_{i=1}^{\ell} (\vec{\nabla}_{r_i} \phi_{i\,\ell+1}) \cdot (\vec{\nabla}_{p_i} f^{(N)}) d^6\vec{x}_{\ell+1} \ldots d^6\vec{x}_N
\end{aligned} \qquad (9.14)$$

since $f^{(N)}$ is symmetric in the $\vec{x}_i$.

Thus,

$$\frac{\partial f^{(\ell)}}{\partial t} = \{H_\ell, f^{(\ell)}\} + \frac{(N-\ell)}{V} \int \sum_{i=1}^{\ell} (\vec{\nabla}_{r_i} \phi_{i\ell+1}) \cdot (\vec{\nabla}_{p_i} f^{(\ell+1)}) d^6\vec{x}_{\ell+1} , \qquad (9.15)$$

where

$$H_\ell = \sum_{i=1}^{\ell} \left(\frac{p_i^2}{2m} + \Phi(\vec{r}_i)\right) + \sum_{1 \leq i < j \leq \ell} \phi_{ij} \quad . \tag{9.16}$$

This set of equations is the BBGKY hierarchy. For the Hamiltonian considered, it is exact and it relates each reduced distribution function to the distribution function of one higher order. These equations are again not useful in this form since, in order to find $f^{(1)}$, $f^{(2)}$ must be known, and so on. Thus, the equation for $f^{(1)}$ implicitly involves all the higher order distribution functions. However, in many physical situations, it is possible to ignore or to approximate the second and higher order distributions and thus to solve for $f^{(1)}$. Also, since it is the case where $N \to \infty$, $V \to \infty$, $\frac{N}{V}$ constant which is of interest, $(N-\ell)/V$ may be replaced by $n \equiv N/V$, slightly simplifying the equations.

§9.2 Vlasov and Boltzmann Equations

As an example, consider the equation for $f^{(1)}$,

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\vec{p}_1}{m} \cdot \vec{\nabla}_{r_1} f^{(1)} - (\vec{\nabla}_{r_1} \Phi(\vec{r}_1)) \cdot (\vec{\nabla}_{p_1} f^{(1)}) = \left.\frac{\partial f^{(1)}}{\partial t}\right|_{\text{coll}} \tag{9.17}$$

where

$$\left.\frac{\partial f^{(1)}}{\partial t}\right|_{\text{coll}} = n \int (\vec{\nabla}_{r_1} \phi_{12}) \cdot (\vec{\nabla}_{p_1} f^{(2)}) d^3\vec{r}_2 d^3\vec{p}_2 \quad . \tag{9.18}$$

If now a system is considered in which the interparticle forces are long-ranged and in which the short-ranged repulsions are unimportant (this may be a good approximation in a dilute plasma), then each particle is influenced by many others at all times. Thus, each particle moves in an averaged "mean field" which is roughly the same for each particle and the particles are not strongly correlated. This makes it possible to write

$$f^{(2)}(\vec{x}_1,\vec{x}_2,t) \approx f^{(1)}(\vec{x}_1,t)f^{(1)}(\vec{x}_2,t) \quad . \qquad (9.19)$$

With this approximation, the equation for $f^{(1)}$ becomes

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\vec{p}_1}{m} \cdot \vec{\nabla}_{r_1} f^{(1)} - (\vec{\nabla}_{r_1}\phi(\vec{r}_1) + \overline{\nabla_{r_1}\phi_{12}}) \cdot \vec{\nabla}_{p_1} f^{(1)} = 0 \qquad (9.20)$$

where $\overline{\nabla_{r_1}\phi_{12}} = n \int (\vec{\nabla}_{r_1}\phi_{12})\, f^{(1)}(\vec{x}_2,t)d^6\vec{x}_2$ (9.21)

is the mean interparticle field gradient.

This equation is known as the Vlasov or mean field equation and is important in plasma physics. It differs from the Boltzmann equation in that the interparticle interactions are treated as a collective potential whereas, in the Boltzmann case, the interaction is assumed to result in discrete binary collisions between almost independent particles.

In order to derive the Boltzmann equation from the BBGKY hierarchy, first consider the original intuitive derivation due to Boltzmann. The most important assumptions in his derivation are that the system is so dilute that only binary collisions occur (this also assumes the forces to be short range), that the effects of external fields upon the dynamics of a collision are negligible, and finally that the number of collisions in a given volume, per unit time, can be statistically computed from $f^{(1)}$ alone. It is this final assumption (Stosszahlansatz) which allows a truncation of the BBGKY hierarchy.

To actually derive the Boltzmann equation from the BBGKY hierarchy, certain assumptions, similar to those made by Boltzmann in his derivation, are necessary. For a dilute system with short range forces (e.g. neutral molecules), the most important feature of the potential, as far as collisions are concerned, is its repulsive core. This core may be approximated as a hard sphere interaction of range σ. If the term

$$\left.\frac{\partial f^{(1)}}{\partial t}\right|_{coll} = n \int (\vec{\nabla}_{r_1}\phi_{12}) \cdot (\vec{\nabla}_{p_1} f^{(2)}) d^3\vec{r}_2 d^3\vec{p}_2 \quad (9.22)$$

is considered for a hard sphere potential, it is seen that $(\vec{\nabla}_{r_1}\phi_{12})$ is zero for $|\vec{r}_1-\vec{r}_2| > \sigma$ and is infinite for $|\vec{r}_1-\vec{r}_2| = \sigma$. Thus, the collision term (9.22) is determined by the value of $f^{(2)}$ near $|\vec{r}_1-\vec{r}_2| = \sigma$. At the instant of collision, the probability that the two hard spheres are contacting a third is infinitesimal so, for $|\vec{r}_1-\vec{r}_2| \approx \sigma$, the function $f^{(2)}$ is just $f^{(2)}$ for a system of only two hard spheres. This is equivalent to Boltzmann ignoring triple collisions in his derivation. Then, $f^{(2)}$ satisfies the equation

$$\frac{\partial f^{(2)}}{\partial t} + \frac{\vec{p}_1}{m} \cdot (\vec{\nabla}_{r_1} f^{(2)}) + \frac{\vec{p}_2}{m} \cdot (\vec{\nabla}_{r_2} f^{(2)}) - (\vec{\nabla}_{r_1}\phi_{12})\cdot(\vec{\nabla}_{p_1} f^{(2)}) - (\vec{\nabla}_{r_2}\phi_{12}) \cdot (\vec{\nabla}_{p_2} f^{(2)}) = 0 \quad . \tag{9.23}$$

If it is assumed that the system has been sufficiently randomized that the entire time dependence of $f^{(2)}$ is through $f^{(1)}$, then $\frac{\partial f^{(2)}}{\partial t} = 0$. This assumption, which allows the truncation of the hierarchy to yield a closed set of equations, is a form of Stosszahlansatz; that is $f^{(2)}$ is a functional of $f^{(1)}$. Then, integrating eqn. (9.23) with respect to $\vec{r}_2$ and $\vec{p}_2$ gives

$$\int \{(\vec{\nabla}_{r_1}\phi_{12})\cdot(\vec{\nabla}_{p_1} f^{(2)}) + (\vec{\nabla}_{r_2}\phi_{12})\cdot(\vec{\nabla}_{p_2} f^{(2)})\} d^3\vec{r}_2 d^3\vec{p}_2$$
$$= \int \{\frac{\vec{p}_1}{m} \cdot (\vec{\nabla}_{r_1} f^{(2)}) + \frac{\vec{p}_2}{m} \cdot (\vec{\nabla}_{r_2} f^{(2)})\} d^3\vec{r}_2 d^3\vec{p}_2 \quad . \tag{9.24}$$

Due to the asymptotic behavior of $f^{(2)}$ as $|\vec{p}_2| \to \infty$, the second term on the left of (9.24) gives zero contribution.

Also, since there are no triple or higher order collisions and the forces are spherically symmetric, $f^{(2)}$ depends only upon $|\vec{r}_1-\vec{r}_2|$. Thus,

$$\vec{\nabla}_{r_1} f^{(2)} = -\vec{\nabla}_{r_2} f^{(2)} \quad . \tag{9.25}$$

Therefore,

$$\begin{aligned} I &\equiv \int (\vec{\nabla}_{r_1}\phi_{12}) \cdot (\vec{\nabla}_{p_1} f^{(2)}) d^3\vec{r}_2 d^3\vec{p}_2 \\ &= \int \frac{(\vec{p}_2-\vec{p}_1)}{m} \cdot (\vec{\nabla}_{r_2} f^{(2)}) d^3\vec{r}_2 d^3\vec{p}_2 \quad . \end{aligned} \tag{9.26}$$

Using spherical coordinates, with $r \equiv |\vec{r}_1-\vec{r}_2|$, gives

$$I = \int \frac{(\vec{p}_2-\vec{p}_1)}{m} \cdot \hat{S} \frac{\partial f^{(2)}}{\partial r} r^2 dr\, d\Omega\, d^3\vec{p}_2 \tag{9.27}$$

where $\hat{S}$ is a unit vector in the direction of $(\vec{r}_2-\vec{r}_1)$ and $d\Omega$ is the differential element of solid angle. Since $\frac{\partial f^{(2)}}{\partial r}$ is non-zero only for $r = \sigma$,

$$I = \sigma^2 \int \frac{(\vec{p}_2-\vec{p}_1)}{m} \cdot \hat{S} f^{(2)} d\Omega\, d^3\vec{p}_2 \quad . \tag{9.28}$$

Next, $f^{(2)}$ is separated into a part $f_1^{(2)}$, representing $f^{(2)}$ before the collision ($f_1^{(2)} = 0$ unless $(\vec{p}_2-\vec{p}_1)\cdot\hat{S} < 0$), and a part $f_2^{(2)}$, representing $f^{(2)}$ after the collision ($f_2^{(2)} = 0$ unless $(\vec{p}_2-\vec{p}_1)\cdot\hat{S} > 0$). Then,

$$I = \sigma^2 \int \frac{|(\vec{p}_2-\vec{p}_1)\cdot\hat{S}|}{m} \{f_2^{(2)} - f_1^{(2)}\} d\Omega\, d^3\vec{p}_2 \quad . \tag{9.29}$$

Finally, since $f^{(2)}$ is the distribution function for a system of only two particles, $f^{(2)}(\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2,t)$ is simply a product of two $f^{(1)}$'s for $|\vec{r}_1-\vec{r}_2| > \sigma$. That is,

$$f_1^{(2)} = f^{(1)}(\vec{r}_1,\vec{p}_1,t)\, f^{(1)}(\vec{r}_2,\vec{p}_2,t)$$

and

$$f_2^{(2)} = f^{(1)}(\vec{r}_1,\vec{p}_1{}',t)\, f^{(1)}(\vec{r}_2,\vec{p}_2{}',t)$$

where $\vec{p}_1$ and $\vec{p}_2$ are the momenta before the collision and $\vec{p}_1{}'$ and $\vec{p}_2{}'$ are the momenta after. If the variation of $f^{(1)}$ is assumed to be small over a distance σ, then, since $|\vec{r}_1-\vec{r}_2| = \sigma$ just before and just after the collision, $\vec{r}_1$ and $\vec{r}_2$ may be replaced by $\vec{r}_1$ in the argument of $f_1^{(1)}$, giving

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\vec{p}_1}{m} \cdot (\vec{\nabla}_{r_1} f^{(1)}) - (\vec{\nabla}_{r_1} \Phi(\vec{r}_1))\cdot(\vec{\nabla}_{p_1} f^{(1)})$$

$$= n\sigma^2 \int \frac{|(\vec{p}_2-\vec{p}_1)\cdot\hat{S}|}{m} \{f^{(1)}(\vec{r}_1,\vec{p}_1{}',t)f^{(1)}(\vec{r}_1,\vec{p}_2{}',t)$$

$$- f^{(1)}(\vec{r}_1,\vec{p}_1,t)f^{(1)}(\vec{r}_1,\vec{p}_2,t)\}\, d\Omega\, d^3p_2 \quad . \quad (9.30)$$

If new distribution functions $f(\vec{r},\vec{v},t) \equiv nm^3 f^{(1)}(\vec{r},\vec{p},t)$ are introduced, eqn. (9.30) becomes

$$\frac{\partial f}{\partial t} + \vec{v}_1 \cdot (\vec{\nabla}_{r_1} f) - \frac{1}{m} (\vec{\nabla}_{r_1}\Phi(\vec{r}_1)) \cdot (\vec{\nabla}_{v_1} f)$$

$$= \sigma^2 \int |(\vec{v}_2-\vec{v}_1)\cdot\hat{S}| \{f(\vec{r}_1,v_1',t)f(\vec{r}_1,\vec{v}_2{}',t)-f(\vec{r}_1,\vec{v}_1,t)$$

$$f(\vec{r}_1,\vec{v}_2,t)\}\, d\Omega\, d^3\vec{v}_2 \quad (9.31)$$

which is the usual form of the Boltzmann equation.

In this equation, the terms on the left represent drifts in real space due to density gradients and drifts in momentum space due to external potentials, while the right side gives the change due to binary collisions.

The range of applicability of the Boltzmann equation is limited by the assumptions which went into its derivation. It obviously does not apply to systems where there are long-range forces or multiple collisions. Also, it is limited to systems without strong field gradients. These restrictions follow from the explicit assumptions which went into the derivation. The restrictions placed upon the system by the use of a Stosszahlansatz are harder to specify. In order to analyse the Stosszahlansatz more carefully, general analysis of the BBGKY hierarchy is required, in the spirit of Bogolyubov [27]. The most general and desirable result would be a systematic method of expanding the distribution functions in powers both of density and of field gradients, in which the statistical assumptions involved are clearly set forth. This is the objective of various methods of formally expanding the BBGKY hierarchy, several of which will be further considered in the next section.

First, a simple example of electrical conductivity will be studied using Boltzmann's equation. Since an attempt to solve the Boltzmann equation directly would involve a detailed description of electron-electron scattering, lattice effects, and impurity structure, the drastic assumption of a constant relaxation time is made right at the start. This assumption may be written as

$$\left.\frac{\partial f^{(1)}}{\partial t}\right|_{coll} = - \frac{(f^{(1)} - f_o^{(1)})}{\tau} \tag{9.32}$$

where τ is a constant relaxation time and $f_o^{(1)}$ is the equilibrium value of $f^{(1)}$ for an ideal gas of conduction electrons in the absence of external fields. This relaxation time approximation is perhaps justified for a system in which the conduction electrons are essentially free (e.g. in a metal) and where the dominant collisions are with impurities and lattice sites (i.e. electron-electron interactions are "smeared out" to give a mean field effect). In this case, the collisions with impurities and with lattice sites are essentially random and a given electron will have a mean free time τ between collisions. If each collision randomizes the direction of motion of a conduction electron and if there are no external fields, the system will return to equilibrium monotonically in a time roughly equal to τ. This may be expressed as

$$(f^{(1)}(t) - f_o^{(1)}) = (f^{(1)}(t=0) - f_o^{(1)})e^{-t/\tau} \tag{9.33}$$

which yields (9.32).

If the system is assumed to be spatially homogeneous ($\vec{\nabla}_{r_1} f^{(1)} = 0$), to be in a steady state ($\frac{\partial f^{(1)}}{\partial t} = 0$), and to be in an external electric field, $\vec{E}$, in the z-direction, then Boltzmann's equation is

$$e\ E_z \frac{\partial f^{(1)}}{\partial p_z} = \frac{(f^{(1)} - f_0^{(1)})}{\tau}$$

or

$$\frac{\partial f^{(1)}}{\partial p_z} - \frac{f^{(1)}}{e\ E_z\ \tau} = - \frac{f_0^{(1)}}{e\ E_z\ \tau} \quad . \tag{9.34}$$

Therefore,

$$f^{(1)} = \int_{p_z}^{\infty} dp_z' \frac{f_0^{(1)}}{e\ E_z\ \tau} \exp\left(- \int_{p_z}^{p_z'} \frac{dp_z''}{e\ E_z\ \tau}\right) \quad . \tag{9.35}$$

For $f_0^{(1)}$ the zero temperature Fermi distribution of an ideal gas,

$$f_0^{(1)} = \theta\ (\varepsilon_F - \varepsilon)$$

where

$$\theta(x) = 1 \quad \text{for} \quad x \geq 0$$

$$\theta(x) = 0 \quad \text{for} \quad x < 0$$

$$\varepsilon = \frac{p^2}{2m} \tag{9.36}$$

and where ε_F is the Fermi energy, it is possible to find $f^{(1)}(t)$ explicitly for zero temperature (cf. Trofimenkoff and Kreuzer [28]). Using their expression for $f^{(1)}$ and the definitions

$$n = \int f^{(1)}\ d^3\vec{p}$$

and

$$j_z = e \int v_z\ f^{(1)}\ d^3p \tag{9.37}$$

for the conduction electron density n and the current j_z yields

$$j_z = (\frac{ne^2\tau}{m})E_z \quad . \tag{9.38}$$

That is, in the relaxation time approximation, Ohm's law holds exactly for all values of E_z with the conductivity given by

$$\sigma = (\frac{ne^2\tau}{m}) \quad . \tag{9.39}$$

This derivation differs from the usual one (e.g. Ziman [29]) in that the Boltzmann equation has not been linearized prior to finding $f^{(1)}$.

§9.3 Formal Methods and Expansions

In order to generalize Boltzmann's kinetic equation, starting from the BBGKY hierarchy, it is useful to define several operators and to examine their properties. Much of the following discussion will follow Bogolyubov [27]. The system considered will again consist of N identical particles with Hamiltonian (9.11).

The distribution function obeys Liouville's equation (9.6). From a formal point of view, the Liouville equation is equivalent to the full dynamical equations of Hamilton. Thus, although $f^{(N)}$ has a statistical interpretation, it is possible in principle to follow the exact time development of $f^{(N)}$. This may be expressed by formally defining an N-particle time development operator $S_t^{(N)}$ (also known as a "streaming operator") which replaces the initial values $\vec{x}_1(o),\ldots,\vec{x}_N(o)$ by their values $\vec{x}_1(t),\ldots,\vec{x}_N(t)$ at time t. That is,

$$f^{(N)}(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N,t) = S_{-t}^{(N)} f^{(N)}(\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_N,o) \quad . \quad (9.40)$$

If also a Liouville operator $\mathcal{L}_N$ is defined by

$$-i\,\mathcal{L}_N\, g = \{g, H_N\} \qquad (9.41)$$

then

$$\frac{\partial f^{(N)}}{\partial t} = i\,\mathcal{L}_N\, f^{(N)}$$

so $f^{(N)}$ may formally be written as

$$f^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,t) = e^{it\mathcal{L}_N}\, f^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,o) \quad . \quad (9.42)$$

Therefore,

$$S_t^{(N)} = e^{-it\mathcal{L}_N} \quad . \qquad (9.43)$$

Also, it is useful to define the auxiliary ℓ-body operators. If $\mathcal{L}_\ell$ and $S_t^{(\ell)}$ are defined by

$$-i\,\mathcal{L}_\ell\, g = \{g, H_\ell\} \qquad (9.44)$$

where H_ℓ was defined in eqn. (9.16), and

$$S_t^{(\ell)} = e^{-it\mathcal{L}_\ell} \qquad (9.45)$$

then the streaming operators have the group property

$$S^{(\ell)}_{t_1}\, S^{(\ell)}_{t_2} = S^{(\ell)}_{t_1+t_2}$$

$$S^{(\ell)}_{o} = 1 \quad . \tag{9.46}$$

Writing the BBGKY hierarchy (9.15) and letting $n = N/V$ and using the Liouville operator (9.41) gives

$$f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,t) = e^{it\mathcal{L}_\ell} f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o) + n\int_o^t e^{i(t-t')\mathcal{L}_\ell} \int\{\sum_{i=1}^{\ell} \phi_{i\ell+1}, f^{(\ell+1)}\}\, d^6\vec{x}_{\ell+1}\, dt$$

$$= e^{it\mathcal{L}_\ell} f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o) + n\int_o^t e^{i(t-t')\mathcal{L}_\ell} \int\{\sum_{i=1}^{\ell} \phi_{i\ell+1}, e^{it'\mathcal{L}_{\ell+1}} f^{(\ell+1)}(\vec{x}_1,\ldots,\vec{x}_{\ell+1},o)\}\, d^6\vec{x}_{\ell+1}\, dt'$$

$$+ n^2\int_o^t e^{i(t-t')\mathcal{L}_\ell} \int\{\sum_{i=1}^{\ell} \phi_{i\ell+1}, \int_o^{t'} e^{i(t'-t'')\mathcal{L}_{\ell+1}} \int\{\sum_{i=1}^{\ell+1} \phi_{i\ell+2}, e^{it''\mathcal{L}_{\ell+2}} f^{(\ell+2)}(\vec{x}_1,\ldots,\vec{x}_{\ell+2},o)\}\, d^6\vec{x}_{\ell+2}\, dt''\}\, d^6\vec{x}_{\ell+1}\, dt'$$

$$+ \ldots \quad . \tag{9.47}$$

The terms in this expansion have simple interpretations. The first term on the right is the unperturbed motion of a system of ℓ particles. The first correction term, proportional to n, describes part of the effect of the other particles. This term results additively from the action of each of the $(N-\ell)$ other particles separately. The term proportional to n^2 results from the interaction of the group of ℓ particles with two other particles, etc.

This expansion, which is equivalent to that derived by Bogolyubov [27], suffers from several disadvantages. Firstly, although it is formally a density expansion, in fact, if the distribution functions were normalized differently, density would not appear in (9.47). Actually, such an expansion must be in powers of some dimensionless parameter $(r_0^3 n)$ where r_0 is a length of order of the range of the interparticle potential. To recast the expansion in terms of such a parameter is not easy, due to the complexity of the terms.

Secondly, it has been found (cf. Cohen [30]) that such an expansion gives rise to secular terms which increase without bound for increasing time and hence any truncation of eqn. (9.47) would lose its validity for large times.

Lastly, this expansion is not useful for computations since, in order to solve for $f^{(\ell)}$, the exact time evolution of $f^{(\ell+1)}, f^{(\ell+2)}, \ldots$ must be known, as in the original hierarchy.

However, since eqn. (9.47) is, in principle, an expansion for the time development of any of the $f^{(\ell)}$, it may be possible to derive less general approximations for the lower order distributions. In particular, it would be useful to generalize Boltzmann's equation for $f^{(1)}$.

In order to do this, it is necessary to make some assumptions regarding relaxation of correlations in the system. This is the vital approximation which allows one to avoid the necessity of solving the complete N-body problem. It also clearly is an assumption which should be carefully analysed before the resulting kinetic equations are applied to a given system.

To justify the approximations to be made, consider the general form of the BBGKY hierarchy

$$\frac{\partial f^{(\ell)}}{\partial t} + \{f^{(\ell)},H_\ell\} = n \int \{\sum_{i=1}^{\ell} \phi_{i\ell+1}, f^{(\ell+1)}\} d^6\vec{x}_{\ell+1} \qquad .(9.48)$$

For $\ell \geq 2$, the expression for $\frac{\partial f^{(\ell)}}{\partial t}$ involves a term in $\{f^{(\ell)},H_\ell\}$ which results from interparticle interactions. In a large system, at any given instant, there will be many collisions occurring and, for $\ell \geq 2$, $f^{(\ell)}$ will change significantly over a time τ_0, of order of the duration of a collision. For the period $t<\tau_0$, the formal expansion (9.47) is useful and a detailed dynamical description is needed. For $t>\tau_0$, the distribution functions for $\ell \geq 2$ have changed drastically but, since $\frac{\partial f^{(1)}}{\partial t}$ does not directly involve the interparticle potential, $f^{(1)}$ has changed much less. Since it is not practical to follow the actual evolution of the $f^{(\ell)}$ for times $> \tau_0$, it is assumed that, in physical systems, correlations disappear for large times. Explicitly, it is assumed that the dynamical evolution of $f^{(\ell)}(t)$ (for $\ell \geq 2$) is completely randomized after the time τ_o, and the only remaining time dependence is through the functional dependence of $f^{(\ell)}$ upon $f^{(1)}$. Thus, after the initial stage ($t<\tau_o$), it is possible to identify a second, "kinetic" stage in which $f^{(1)}$ alone is sufficient to describe the system. In this stage, all the initial correlations have died out and the system can be described statistically in terms of $f^{(1)}$. This is the generalization of the Stosszahlansatz. The system also possesses a second characteristic time $t_c>\tau_o$, of order of a particle's mean time between collisions. For $t>>t_c$, nearly all particles have undergone collisions and these collisions have changed $f^{(1)}$ to an almost equilibrium distribution. In accordance with the ideas of local equilibrium in Chapter I, for $t>>t_c$, it is no longer necessary to know $f^{(1)}$ in order to describe the system. In this third, "hydrodynamical" stage, the macroscopic mass, energy, and momentum densities are sufficient.

It should be noted that, if a system has long range forces, or is very dense, the time τ_o may be of the same magnitude as t_c, and no kinetic stage may be identified. However, collective interactions may still return the system to equilibrium very rapidly.

To derive kinetic equations, it is first assumed that the correlation functions $f^{(\ell)}$ for $t>0$ are determined by the value of $f^{(1)}$ at $t=0$. That is,

$$f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,t) = f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell;f^{(1)}) \quad . \qquad (9.49)$$

Such a solution is not the most general, since the initial distribution $f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o)$ cannot be specified arbitrarily with this ansatz, but is fully determined by $f^{(1)}(\vec{x}_1,o)$, thus incorrectly eliminating some physically realizable initial conditions. However, in view of the above arguments, it is hoped that by "coarse-graining" (i.e. averaging) over a time of order τ_o, the necessity of considering the precise initial conditions can be eliminated.

In order to analytically formulate the condition that the correlations may be ignored, define an operator $G_t^{(\ell)}$ and a correlation function $a^{(\ell)}$ by

$$f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o)=a^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o)\prod_{i=1}^{\ell} f^{(1)}(\vec{x}_i,o) \qquad (9.50)$$

and

$$G_t^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell)=S_{-t}^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell)a^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o)\prod_{i=1}^{\ell} S_t^{(1)}(\vec{x}_i) \quad . \qquad (9.51)$$

Then, $a^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,o)=1$ if there are no correlations at $t=0$ and

$$f^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell,t) = G_t^{(\ell)}(\vec{x}_1,\ldots,\vec{x}_\ell)\prod_{i=1}^{\ell} f^{(1)}(\vec{x}_i,t) \ . \qquad (9.52)$$

The assumption that there are no second order correlations present at time t may be written as

$$G_t^{(2)}(\vec{x}_1,\vec{x}_2) = S_{-t}^{(2)}(\vec{x}_1,\vec{x}_2)\ S_t^{(1)}(\vec{x}_1)\ S_t^{(1)}(\vec{x}_2) \qquad (9.53)$$

obtained by setting $a^{(2)} = 1$. This is an assumption which implies that two states with the same values of $f^{(1)}$ will evolve to the same state at a time $t >> \tau_o$, regardless of initial correlations. This can only be true in a coarse grained sense, but this may be sufficient. For $|\vec{r}_1-\vec{r}_2| >> r_o$, $G_t^{(2)}(\vec{x}_1,\vec{x}_2)$ is independent of t since $S_{-t}^{(2)}(\vec{x}_1,\vec{x}_2)$ differs from $S_{-t}^{(1)}(\vec{x}_1)\ S_{-t}^{(1)}(\vec{x}_2)$ only if the particles are separated by less than r_o. Writing

$$G_\infty^{(2)} \equiv \lim_{t\to\infty} G_t^{(2)}$$

and ignoring correlations, the BBGKY equation for $f^{(1)}$, in the kinetic stage, is

$$\frac{\partial f^{(1)}}{\partial t} + \{f^{(1)},H_1\} = n\int\{\phi_{12},G_\infty^{(2)}(\vec{x}_1,\vec{x}_2)f^{(1)}(\vec{x}_1,t)f^{(1)}(\vec{x}_2,t)\}\, d^6\vec{x}_2 \ . \qquad (9.54)$$

Using the notation

$$G_\infty^{(2)}(\vec{x}_1,\vec{x}_2)\vec{x}_i = X_i^{(2)}(\vec{x}_1,\vec{x}_2) = (\vec{R}_i^{(2)}(\vec{x}_1,\vec{x}_2),\vec{P}_i^{(2)}(\vec{x}_1,\vec{x}_2)) \qquad (9.55)$$

to denote the initial position and momentum of the ith particle and defining $\vec{r}_{21} \equiv \vec{r}_2 - \vec{r}_1$ gives

$$\vec{R}_1^{(2)}(\vec{x}_1,\vec{x}_2) = \vec{r}_1 + \frac{1}{2}\vec{r}_{21} - \frac{1}{2}G_\infty^{(2)}(\vec{x}_1,\vec{x}_2)\vec{r}_{21}$$

$$\vec{R}_2^{(2)}(\vec{x}_1,\vec{x}_2) = \vec{r}_1 + \frac{1}{2}\vec{r}_{21} + \frac{1}{2}G_\infty^{(2)}(\vec{x}_1,\vec{x}_2)\vec{r}_{21} \quad . \qquad (9.56)$$

Thus,

$$n\int\{\phi_{12},G_\infty^{(2)}(\vec{x}_1,\vec{x}_2)f^{(1)}(\vec{x}_1,t)f^{(1)}(\vec{x}_2,t)\}d^6\vec{x}_2$$

$$= n\,[J(f^{(1)},f^{(1)}) + J_1(f^{(1)},f^{(1)})] \qquad (9.57)$$

where

$$J(f^{(1)},f^{(1)}) = \int\{\phi_{12},f^{(1)}(\vec{r}_1,\vec{p}_1^{(2)},t)f^{(1)}(\vec{r}_1,\vec{p}_2^{(2)},t)\}d^6\vec{x}_2 \qquad (9.58)$$

and where $J_1(f^{(1)},f^{(1)})$ contains correction terms which may be ignored if the change in $f^{(1)}$ is small over a distance of the order of the interaction range. The term involving $J(f^{(1)},f^{(1)})$ is identical to that considered in §9.2 and, by the same arguments, can be reduced to the usual form of the Boltzmann collision term. The correction term involving J_1 can be calculated for a hard sphere potential and is found to be important only for strongly inhomogeneous distribution functions.

The extension to higher densities by including some of the correlations is quite involved. However, the Boltzmann equation is sufficient to describe moderately dense gases and, in order to describe liquids or solids, many terms would have to be kept, since there are strong correlations present. Therefore, it may be that such extensions are not very important in a general survey of non-equilibrium systems and are of limited calculational benefit due to their complexity.

In the next section, balance equations for continuous media will be derived directly from the BBGKY hierarchy. Before presenting the derivation of these equations, it will be shown how the methods used in analysing the BBGKY hierarchy can be modified to deal with the response of a system to an external perturbation of the Hamiltonian.

Consider an isolated system with Hamiltonian H_o. If an external perturbation, described by an interaction Hamiltonian H_t', is switched on at t=0, then the perturbed and unperturbed motions of the system will be different. If $f_o^{(N)}$ is the distribution function of the unperturbed system and if $f^{(N)} = f_o^{(N)} + f^{(N)'}$ is the distribution function of the perturbed system, then

$$\frac{\partial f_o^{(N)}}{\partial t} + \{f_o^{(N)}, H_o\} = 0 \qquad (9.59)$$

and

$$\frac{\partial f^{(N)}}{\partial t} + \{f^{(N)}, H_o + H_t'\} = 0 \quad . \qquad (9.60)$$

Eqn. (9.60) may be written as

$$\frac{\partial f^{(N)}}{\partial t} + \{f^{(N)},H_o\} = - \{f^{(N)},H_t'\} \quad . \qquad (9.61)$$

Using the Liouville operator $\mathcal{L}_N$, defined for H_o, this can be converted to an integral equation as

$$f^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,t)=e^{it\mathcal{L}_N}f_o^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,o)-\int_o^t e^{i(t-t')\mathcal{L}_N}$$

$$\{f^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,t),H_t'\}dt' \quad . \qquad (9.62)$$

If $e^{it\mathcal{L}_N}f_o(\vec{x}_1,\ldots,\vec{x}_N,o)$ is written as $f_o^{(N)}(\vec{x}_1,\ldots,\vec{x}_N,t)$ and the interaction Hamiltonian is written as

$$H_t' = - \mu\ F(t)\ A \qquad (9.63)$$

where A is a time independent operator, F(t) is a function of time which is zero for t<0, and μ is a coupling constant, then iterating eqn. (9.62) yields

$$f^{(N)}(t)=f_o^{(N)}(t)+\sum_{k=1}^{\infty}(\mu)^k\int_o^t dt'\int_o^{t'} dt''\ldots\int_o^{t^{(k-1)}} dt^{(k)}[e^{i(t-t')\mathcal{L}_N}$$

$$x\{e^{i(t'-t'')\mathcal{L}_N},\{\ldots,\{e^{i(t^{(k-1)}-t^{(k)})\mathcal{L}_N}\{f_o^{(N)}(t^{(k)}),$$

$$A\}\ldots\}F(t')\ldots F(t^{(k)})] \quad . \qquad (9.64)$$

For a time dependent $f^{(N)}$, eqn. (9.63) suffers from the same drawback that the expansion (9.47) of the BBGKY hierarchy had, namely that, as t increases, the accuracy of a truncation of the series decreases. Thus, for general time-dependent problems, the validity of a finite sum of terms in eqn. (9.64) is limited to very short times. The time range in which such an expansion can be valid will be investigated in more detail in Chapter XI.

On the other hand, steady state solutions long after the perturbation has been switched on are often of interest, particularly for transport phenomena. In this case, it is more convenient to assume that the system is in equilibrium at $t=-\infty$ and that H_t' is switched on adiabatically, starting at $t=-\infty$. That is, assume

$$f_0^{(N)}(\vec{x}_1,\dots,\vec{x}_N,t)=f_0^{(N)}(\vec{x}_1,\dots,\vec{x}_N,t=-\infty)\alpha\, e^{-\beta H_0} \quad . \qquad (9.65)$$

In this case, the expansion for $f^{(N)}(t)$ is

$$f^{(N)}(t) = f_0^{(N)} + \sum_{k=1}^{\infty} (\mu)^k \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \dots \int_{-\infty}^{t^{(k-1)}} dt^{(k)}$$

$$[e^{i(t-t')\mathcal{L}_N} \times \{e^{i(t'-t'')\mathcal{L}_N} \dots, \{e^{i(t^{(k-1)}-t^{(k)})\mathcal{L}_N},$$

$$\{f_0^{(N)},A\}\dots\}F(t')\dots F(t^{(k)})] \quad . \qquad (9.66)$$

This equation again requires an exact knowledge of $e^{it\mathcal{L}_N}$, the time development operator of the unperturbed system, for its solution. Its quantum mechanical analogue, the response theory of Kubo, is much studied and, for example, is used to derive Onsager relations, fluctuation-dissipation theorems, and sum rules in the steady state.

§9.4 Microscopic Derivation of Balance Equations

In order to derive macroscopic balance equations from microscopic foundations, it is necessary to express the macroscopic densities in terms of the distribution functions. The macroscopic mass density $\rho(\vec{r},t)$, fluid velocity $\vec{v}(\vec{r},t)$, thermal kinetic energy density $\frac{3}{2}\,\frac{k}{m}\,T(\vec{r},t)$, and mean interparticle potential energy density $\varepsilon_\phi(\vec{r},t)$, may be defined as

$$\rho(\vec{r},t) = m\left(\frac{N}{V}\right) \int f^{(1)}(\vec{r},\vec{p},t)\,d^3\vec{p} \tag{9.67}$$

$$\vec{v}(\vec{r},t) = \frac{1}{\rho}\left(\frac{N}{V}\right) \int \vec{p}\,f^{(1)}(\vec{r},\vec{p},t)\,d^3\vec{p} \tag{9.68}$$

$$\frac{3}{2}\,\frac{k}{m}T(\vec{r},t) = \frac{1}{\rho}\left(\frac{N}{V}\right) \int \frac{(\vec{p}-m\vec{v})^2}{2m} f^{(1)}(\vec{r},\vec{p},t)\,d^3\vec{p} \tag{9.69}$$

and

$$\varepsilon_\phi(\vec{r},t) = \frac{1}{\rho}\left(\frac{N}{V}\right)^2 \int \phi(|r-r_1|)\,f^{(2)}(\vec{r},\vec{p},\vec{r}_1,\vec{p}_1,t)\,d^3\vec{p}\,d^3\vec{r}_1\,d^3\vec{p}_1 \quad , \tag{9.70}$$

in agreement with the usual microscopic interpretation of these quantities. As in Chapter II, the energy densities are specific densities.

To derive balance equations for the above quantities, consider the first BBGKY equation,

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\vec{p}}{m} \cdot (\vec{\nabla}_r f^{(1)}) + \vec{F}\cdot(\vec{\nabla}_p f^{(1)}) = \left(\frac{N}{V}\right)\int (\vec{\nabla}_r \phi(|\vec{r}-\vec{r}_1|)\cdot(\vec{\nabla}_p f^{(2)})\; d^3\vec{r}_1\,d^3\vec{p}_1 \tag{9.71}$$

where $-\nabla_r\Phi(\vec{r})$ has been replaced by $\vec{F}$ to conform to the notation of Chapter II. First, simply integrate this equation with respect to $\vec{p}$. This gives

$$\int\frac{\partial f^{(1)}}{\partial t}d^3\vec{p} + \int\frac{\vec{p}}{m}\cdot\vec{\nabla}_r f^{(1)}d^3\vec{p} - \left(\frac{N}{V}\right)\int(\vec{\nabla}_r\phi)\cdot(\vec{\nabla}_p f^{(2)})d^3\vec{r}_1 d^3\vec{p}_1 d^3\vec{p}=0 \ . \tag{9.72}$$

This may be written as

$$\frac{\partial}{\partial t}[m\left(\frac{N}{V}\right)\int f^{(1)}d^3\vec{p}] + \vec{\nabla}_r\cdot[m\left(\frac{N}{V}\right)\int\frac{\vec{p}}{m}f^{(1)}d^3\vec{p}]$$
$$+ m\left(\frac{N}{V}\right)\int\vec{\nabla}_p\cdot(f^{(1)}\vec{F})d^3\vec{p}-m\left(\frac{N}{V}\right)^2\int\vec{\nabla}_p\cdot[(\vec{\nabla}_r\phi)f^{(2)}]d^3\vec{r}_1 d^3\vec{p}_1 d^3\vec{p}=0 \ . \tag{9.73}$$

Using the divergence theorem and the asymptotic behavior of $f^{(1)}$ and $f^{(2)}$ as $|p|\to\infty$ shows that the last two integrals are zero, so

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot(\rho\vec{v}) = 0 \quad . \tag{9.74}$$

This is just the continuity equation for a macroscopic medium, derived from microscopic foundations.

For the second balance equation, multiply eqn. (9.71) by $\vec{p}$ and then integrate with respect to $\vec{p}$. This yields

$$\frac{\partial}{\partial t}\left(\frac{N}{V}\right)\int f^{(1)}\vec{p}d^3\vec{p}+\vec{\nabla}_r\cdot\left(\frac{N}{V}\right)\int\frac{\vec{p}\vec{p}}{m}f^{(1)}d^3\vec{p}+\left(\frac{N}{V}\right)\int\vec{\nabla}_p\cdot(\vec{p}\vec{F}f^{(1)})d^3\vec{p}$$
$$-\left(\frac{N}{V}\right)\int f^{(1)}\vec{F}\cdot(\vec{\nabla}_p\vec{p})d^3\vec{p}-\left(\frac{N}{V}\right)^2\int\vec{p}(\vec{\nabla}_r\phi)\cdot(\vec{\nabla}_p f^{(2)})d^3\vec{r}_1 d^3\vec{p}_1 d^3\vec{p}=0 \tag{9.75}$$

or

$$\frac{\partial(\rho\vec{v})}{\partial t} + \vec{\nabla}_r \cdot \left(\frac{N}{V}\right) \int \frac{\vec{p}\vec{p}}{m} f^{(1)} \, d^3\vec{p}$$

$$- \left(\frac{N}{V}\right)^2 \int \vec{p}(\vec{\nabla}_r\phi)\cdot(\vec{\nabla}_p f^{(2)}) d^3\vec{r}_1 d^3\vec{p}_1 d^3\vec{p} = \rho\vec{F} \qquad (9.76)$$

where the divergence theorem was used again to show that

$$\int \vec{\nabla}_p \cdot (\vec{p}\vec{F} \, f^{(1)}) d^3\vec{p} = 0 \quad . \qquad (9.77)$$

Next, multiply eqn. (9.74) by $\vec{v}$ to get

$$\vec{v}\,\frac{\partial\rho}{\partial t} + \vec{\nabla}_r \cdot (\rho\vec{v}\vec{v}) - \rho(\vec{v}\cdot\vec{\nabla}_r)\vec{v} = 0 \quad . \qquad (9.78)$$

Subtracting eqn. (9.78) from eqn. (9.76) gives

$$\rho\left(\frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\vec{\nabla}_r)\vec{v}\right) + \vec{\nabla}_r\cdot\left[\left(\frac{N}{V}\right)\int f^{(1)}\left(\frac{\vec{p}\vec{p}}{m} - m\vec{v}\vec{v}\right)d^3\vec{p}\right.$$

$$+ \left(\frac{N}{V}\right)^2 \int \rho^{(2)}(\vec{\nabla}_r\phi) d^3\vec{r}_1 = \rho\vec{F} \qquad (9.79)$$

where, for convenience, $\rho^{(2)}$ has been defined as

$$\rho^{(2)}(\vec{r}_1, \vec{r}_2, t) = \int f^{(2)}(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) d^3\vec{p}_1 d^3\vec{p}_2 \quad . \qquad (9.80)$$

Using the definition of $\vec{v}$,

$$\vec{\nabla}_r\cdot\left[\left(\frac{N}{V}\right)\int f^{(1)}\left(\frac{\vec{p}\vec{p}}{m} - m\vec{v}\vec{v}\right)d^3\vec{p}\right] = \vec{\nabla}_r\cdot\left[\left(\frac{N}{V}\right)\int f^{(1)} \, \frac{(\vec{p}-m\vec{v})(\vec{p}-m\vec{v})}{m} d^3\vec{p}\right] \quad . \qquad (9.81)$$

Thus, defining a kinetic pressure tensor $\overleftrightarrow{P}^{(k)}$ by

$$\overleftrightarrow{P}^{(k)}_{\alpha\beta} = \left[\left(\frac{N}{V}\right) \int f^{(1)} \, \frac{(\vec{p}_\alpha - m\vec{v}_\alpha)(\vec{p}_\beta - m\vec{v}_\beta)}{m} \, d^3\vec{p}\right] \qquad (9.82)$$

gives the balance equation as

$$\rho\left(\frac{\partial\vec{v}}{\partial t}+(\vec{v}\cdot\vec{\nabla}_r)\vec{v}\right) = \rho\vec{F}-\vec{\nabla}_r\cdot\overleftrightarrow{P}^{(k)}-\left(\frac{N}{V}\right)^2\int\rho^{(2)}(\vec{\nabla}_r\phi)\,d^3\vec{r}_1 \quad . \qquad (9.83)$$

In order to cast the last expression as the divergence of a pressure tensor, write the functions in terms of the center of mass and relative coordinates

$$\vec{R} = \frac{\vec{r}+\vec{r}_1}{2}$$

and

$$\vec{r}\,' = \vec{r}_1-\vec{r} \quad . \qquad (9.84)$$

Then,

$$\vec{\nabla}_r\phi = -\vec{\nabla}_{r'}\phi = -\frac{\partial\phi(r')}{\partial r'}\frac{\vec{r}\,'}{r'} \qquad (9.85)$$

where $r' = |\vec{r}\,'|$. Then, rewrite $\rho^{(2)}$ in center of mass coordinates as $\bar{\rho}^{(2)}$, defined by

$$\bar{\rho}^{(2)}(\vec{r}\,',\vec{R},t) = \rho^{(2)}\left(\vec{R}-\frac{\vec{r}\,'}{2},\vec{R}+\frac{\vec{r}\,'}{2},t\right) \quad . \qquad (9.86)$$

The last term in eqn. (9.82) becomes

$$-\left(\frac{N}{V}\right)^2\int\bar{\rho}^{(2)}(\vec{r}\,',\vec{R},t)\frac{\vec{r}\,'}{r'}\frac{\partial\phi}{\partial r'}\,d^3\vec{r}_1 \quad . \qquad (9.87)$$

To write this as a divergence, expand $\bar{\rho}^{(2)}$ about $\vec{r}$ as

$$\begin{aligned}\bar{\rho}^{(2)}(\vec{r}\,',\vec{R},t) &= \bar{\rho}^{(2)}(\vec{r}\,',\vec{r},t) + (\vec{R}-\vec{r})\cdot\vec{\nabla}_r\bar{\rho}^{(2)}(\vec{r}\,',\vec{r},t) + \ldots \\ &= \bar{\rho}^{(2)}(\vec{r}\,',\vec{r},t) + \frac{\vec{r}\,'}{2}\cdot\vec{\nabla}_r\bar{\rho}^{(2)}(\vec{r}\,',\vec{r},t) + \ldots \quad .\end{aligned} \qquad (9.88)$$

Since the integral is over regions where $\frac{\partial\phi}{\partial r'} \neq 0$ (i.e. over the range of the interparticle potential), if the distribution function $\bar{\rho}^{(2)}$ varies slowly with $\vec{R}$ over molecular dimensions, then only the first two terms in the expansion need be kept. Note that, since $\rho^{(2)}$ is symmetric in $\vec{r}$ and $\vec{r}_1$, $\bar{\rho}^{(2)}$ is even in $\vec{r}'$. Thus, expanding $\bar{\rho}^{(2)}$ in expression (9.87) and changing the integration variable from $\vec{r}_1$ to $\vec{r}'$ gives

$$(\frac{N}{V})^2\int\bar{\rho}^{(2)}(\vec{r}',\vec{R},t)\frac{\vec{r}'}{r'}\frac{\partial\phi}{\partial r'}\,d^3\vec{r}_1 \approx (\frac{N}{V})^2\int\bar{\rho}^{(2)}(\vec{r}',\vec{r},t)\frac{\vec{r}'}{r'}\frac{\partial\phi}{\partial r'}\,d^3\vec{r}'$$

$$+ (\frac{N}{V})^2\,\vec{\nabla}_r\cdot\int\frac{r'}{2}\bar{\rho}^{(2)}(\vec{r}',\vec{r},t)\frac{\vec{r}'}{r'}\frac{\partial\phi}{\partial r'}\,d^3\vec{r}'$$

$$= (\frac{N}{V})^2\vec{\nabla}_r\cdot\int\frac{\vec{r}'\vec{r}'}{2r'}\frac{\partial\phi}{\partial r'}\,\bar{\rho}^{(2)}(\vec{r}',\vec{r},t)\,d^3\vec{r}' \quad . \qquad (9.89)$$

Therefore, defining an intermolecular pressure tensor $\overleftrightarrow{P}^{(\phi)}$ by

$$\overleftrightarrow{P}^{(\phi)}_{\alpha\beta} = \frac{1}{2}(\frac{N}{V})^2\int\frac{\vec{r}'_\alpha\vec{r}'_\beta}{r'}\;\frac{\partial\phi}{\partial r'}\,\bar{\rho}^{(2)}(\vec{r}',\vec{r},t)d^3\vec{r}' \qquad (9.90)$$

gives the balance equation in the familiar form

$$\rho(\frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\vec{\nabla}_r)\vec{v}) = \rho\vec{F} - \vec{\nabla}_r\cdot\overleftrightarrow{P} \qquad (9.91)$$

where the total pressure $\overleftrightarrow{P}$ is just $(\overleftrightarrow{P}^{(k)} + \overleftrightarrow{P}^{(\phi)})$.

As implied by its name, $\overleftrightarrow{P}^{(k)}$ is a purely kinetic term and would, for example, be present in an ideal gas. The term $\overleftrightarrow{P}^{(\phi)}$ is due to the force exerted on a fluid element due to the short range intermolecular forces.

The balance equation is exact to the extent that variations of $\bar{\rho}^{(2)}$ are small when $\vec{R}$ varies over distances of the order of the interparticle forces. This would be satisfied in the kinetic and hydrodynamical stages in a system. In a situation where this balance equation fails, any generalization of the pressure tensor would be of little practical value, since the rapid spatial variations in $\bar{\rho}^{(2)}$ would mean that a microscopic description is necessary. Thus, the given expression for the pressure tensor is valid in the domain where a macroscopic continuum description of the system is sufficient.

From the expression for $\overleftrightarrow{P}$, the hydrostatic pressure p may also be found as

$$p \equiv \frac{1}{3}\mathrm{Tr}\overleftrightarrow{P} = \frac{1}{3}\left(\frac{N}{V}\right) \int \frac{(\vec{p}-m\vec{v})^2}{m} f^{(1)} d^3\vec{p} + \frac{1}{6}\left(\frac{N}{V}\right)^2 \int \bar{\rho}^{(2)} r' \frac{\partial\phi}{\partial r'} d^3\vec{r}' \quad . \qquad (9.92)$$

Using the definition (9.69), this may be written

$$p = \rho\frac{kT(\vec{r},t)}{m} + \frac{1}{6}\left(\frac{N}{V}\right)^2 \int \bar{\rho}^{(2)} r' \frac{\partial\phi}{\partial r'} d^3\vec{r}' \quad . \qquad (9.93)$$

The first term in eqn. (9.93) is just the ideal gas pressure and the second is the correction due to interactions. The interaction term varies in rather a complicated manner, but at low densities is proportional to n^2 and decreases with increasing temperature. In a gas, where n is small, the kinetic pressure term dominates, but in liquids, the second term is important and dominates the behavior of p. This is the basis of the qualitative differences between pressures (and similarly viscosities) of liquids and gases.

In a similar manner, it is possible to derive a balance equation for the total internal energy density $u(\equiv \frac{3}{2}\frac{kT}{m} + \varepsilon_\phi)$:

$$\rho(\frac{\partial u}{\partial t} + (\vec{v}\cdot\vec{\nabla}_r)u) + \vec{\nabla}_r\cdot\vec{J}_q = \overleftrightarrow{P} : (\vec{\nabla}_r\vec{v}) \qquad (9.94)$$

where

$$\vec{J}_q = \vec{J}_q^{(k)} + \vec{J}_q^{(\phi_1)} + \vec{J}_q^{(\phi_2)} \qquad (9.95)$$

and

$$\vec{J}_q^{(k)} = (\frac{N}{V})\int \frac{\vec{p}}{m}\frac{(\vec{p}-m\vec{v})^2}{m} f^{(1)}(\vec{r},\vec{p},t)d^3\vec{p} \quad , \qquad (9.96)$$

$$\vec{J}_q^{(\phi_1)} = \frac{1}{2}(\frac{N}{V})^2\int\frac{(\vec{p}-m\vec{v})}{m}\phi(|\vec{r}-\vec{r}_1|)f^{(2)}(\vec{r},\vec{p},\vec{r}_1,\vec{p}_1,t)d^3\vec{r}_1d^3\vec{p}_1d^3\vec{p} \quad , \qquad (9.97)$$

and

$$\vec{J}_q^{(\phi_2)} = \frac{1}{4}(\frac{N}{V})^2\int\frac{\partial\phi}{\partial r'}\frac{\vec{r}'}{r'}(\frac{\vec{r}'}{m})\cdot[(\vec{p}-m\vec{v})+(\vec{p}_1-m\vec{v})]\overline{f}^{(2)}(\vec{r}',\vec{r};\vec{p},\vec{p}_1,t)$$
$$d^3\vec{r}_1d^3\vec{p}_1d^3\vec{p} \quad . \qquad (9.98)$$

In the expression for $\vec{J}_q^{(\phi_2)}$, $\overline{f}^{(2)}(\vec{r}',\vec{r};\vec{p},\vec{p}_1,t)$ is the second order distribution function defined in terms of the center of mass and relative coordinates by

$$\overline{f}^{(2)}(\vec{r}',\vec{R};\vec{p}_1,\vec{p},t) = f^{(2)}(\vec{R}-\frac{\vec{r}'}{2},\vec{p},\vec{R}+\frac{\vec{r}'}{2},\vec{p}_1,t) \quad . \qquad (9.99)$$

For details of the derivation see, for example, Green [31].

With the above definitions of P and $\vec{J}_q$, the balance equations are precisely in the form of Chapter II.

Moreover, they give microscopic formulae for the phenomenological quantities $\overleftrightarrow{P}$ and $\vec{J}_q$. Although $f^{(1)}$ and $f^{(2)}$ are not precisely known, using suitable approximations to them can yield useful results.

Before leaving this section, note that the balance equations can be generalized to multicomponent systems. Also, note that $\overleftrightarrow{P}_{\alpha\beta} = \overleftrightarrow{P}_{\beta\alpha}$. This symmetry, which is assumed in classical hydrodynamics and which was discussed in §2.1, is rigorously valid only for central two-body interactions, since this assumption is implicit in the form of the BBGKY hierarchy.

The various methods and results of this chapter are quite easily generalized to the quantum case, and this will be done in some detail in later chapters. First, in Chapter X, a discussion will be given of some of the fundamental problems of irreversibility itself in classical systems.

CHAPTER X

IRREVERSIBILITY AND *H*-THEOREMS

§10.1 Boltzmann's Equation and Irreversibility

Classical thermodynamics was based almost entirely upon the idea of equilibrium states. As such, it was a logically complete theory which did not concern itself with non-equilibrium situations. However, it seemed desirable to supplement the theory with a statement that all physical systems either were in equilibrium or would reach equilibrium if given enough time. Such a statement would be in agreement with many observations, such as the tendency of a gas to expand to fill a container or for a solid to reach a uniform temperature equal to that of its surroundings. A statement of this irreversible tendency to equilibrium was supplied by the second law of thermodynamics which states that entropy always increases or, at best, in equilibrium remains constant. As long as thermodynamics was regarded as a fundamental theory, such an approach was satisfactory.

However, with the advent of kinetic theory and statistical mechanics, an explanation of observed irreversible behavior, in terms of a microscopic theory, became necessary. The first serious attempt to provide such an explanation was due to Boltzmann and was expressed in his famous *H*-theorem.

In order to demonstrate the irreversibility inherent in the Boltzmann equation, consider the case where there are no external fields ($\Phi(\vec{r}) = 0$) and the distribution function is initially homogeneous ($\vec{\nabla}_r f(\vec{r},\vec{v},t) = 0$). The Boltzmann equation is then

$$\frac{\partial f_1}{\partial t} = \sigma^2 \int |(\vec{v}_2-\vec{v}_1)\cdot\hat{S}|\{f_1'f_2'-f_1f_2\}d\Omega d\vec{v}_2 \qquad (10.1)$$

where, for convenience, f_1, f_2, f_1', and f_2' have been defined by

$$f_1 = f(\vec{r}_1,\vec{v}_1,t) \;;\; f_2 = f(\vec{r}_1,\vec{v}_2,t)$$

$$f_1' = f(\vec{r}_1,\vec{v}_1',t) \;;\; f_2' = f(\vec{r}_1,\vec{v}_2',t) \qquad (10.2)$$

and, as in §9.2, f is the single particle distribution function, the primes denote velocities before a binary collision, and the unprimed velocities are after the collision.

Then, defining an H-quantity by

$$H = \int f_1 \ln f_1 \, d^3\vec{v}_1 d^3\vec{r}_1 \quad , \qquad (10.3)$$

$\frac{dH}{dt}$ is found to be

$$\frac{dH}{dt} = \int \frac{\partial f_1}{\partial t} [1 + \ln f_1] d^3\vec{v}_1 d^3\vec{r}_1 \qquad (10.4)$$

and, since

$$\int \frac{\partial f_1}{\partial t}\, d^3\vec{v}_1 d^3\vec{r}_1 = \frac{d}{dt}\int f_1 d^3\vec{v}_1 d^3\vec{r}_1 = \frac{dN}{dt} = 0 \qquad (10.5)$$

for an isolated system, using Boltzmann's equation, eqn. (10.4) may be rewritten as

$$\frac{dH}{dt} = \int \ln f_1 \{\sigma^2 \int |(\vec{v}_2-\vec{v}_1)\cdot\hat{S}|(f_1'f_2'-f_1f_2)\,d\Omega\, d^3\vec{v}_2\}\, d^3\vec{v}_1 d^3\vec{r}_1 \;. \qquad (10.6)$$

From the form of $\frac{dH}{dt}$, it is apparent that $\frac{dH}{dt}$ is invariant under exchange of $\vec{v}_1$ and $\vec{v}_2$, so

$$\frac{dH}{dt} = \frac{1}{2}\int \ln(f_1f_2)\{\sigma^2 \int |(\vec{v}_2-\vec{v}_1)\cdot\hat{S}|(f_1'f_2'-f_1f_2)\,d\Omega\, d^3\vec{v}_2\}\, d^3\vec{v}_1 d^3\vec{r}_1 \;. \qquad (10.7)$$

Also, from the symmetry between direct and inverse collisions (cf. Terletskii [32] or Kestin and Dorfman [33]), $\frac{dH}{dt}$ is invariant under exchange of $\{\vec{v}_1,\vec{v}_2\}$ and $\{\vec{v}_1',\vec{v}_2'\}$. That is,

$$\frac{dH}{dt} = \frac{1}{2}\int \ln(f_1'f_2')\{\sigma^2 \int |(\vec{v}_2'-\vec{v}_1')\cdot\hat{S}|(f_1f_2-f_1'f_2')\,d\Omega\, d^3\vec{v}_2'\}\, d^3\vec{v}_1' d^3\vec{r}_1 \;. \qquad (10.8)$$

Combining (10.7) and (10.8) (using the facts that $d^3\vec{v}_2' d^3\vec{v}_1' = d^3\vec{v}_2 d^3\vec{v}_1$ and $|(\vec{v}_2'-\vec{v}_1')\cdot\hat{S}| = |(\vec{v}_2-\vec{v}_1)\cdot\hat{S}|$) gives

$$\frac{dH}{dt} = \frac{1}{4}\iint \sigma^2 |(\vec{v}_2-\vec{v}_1)\cdot\hat{S}|(f_1'f_2'-f_1f_2)\ln\left(\frac{f_1f_2}{f_1'f_2'}\right)d\Omega\, d^3\vec{v}_2 d^3\vec{v}_1 d^3\vec{r}_1 \;. \qquad (10.9)$$

Due to the purely mathematical relationship

$$(x-y)\ln\left(\frac{x}{y}\right) \geq 0 \qquad (10.10)$$

for all x and y, it is seen that

$$\frac{dH}{dt} \leq 0 \qquad (10.11)$$

for all times and that the equality holds only for $f_1 f_2 = f_1' f_2'$. This demonstrates that the Boltzmann equation describes irreversible evolution of the single particle distribution function from an arbitrary initial distribution to a final state in which $f_1 f_2 = f_1' f_2'$.

When Boltzmann presented his *H*-theorem, the atomic nature of matter and the kinetic theory of gases were not well-established and his theory met strong opposition. Two of the most important objections to the *H*-theorem were in the form of paradoxes.

The first "paradox", due to Loschmidt (1876), may be stated as follows. The basic laws of mechanics, which govern any isolated system, are time reversal invariant in the sense that, for any possible motion of the system, instantaneously reversing the velocities of all the particles results in the system retracing its trajectory back to its initial state. Thus, for any evolution of a system towards equilibrium, there is another possible motion which takes the system away from equilibrium. Therefore, there seems to be a direct contradiction between the microscopic reversibility of the dynamics of the system and the proven irreversibility of the Boltzmann equation.

The second "paradox", raised by Zermelo, is based on the Poincaré recurrence theorem. This theorem, loosely stated, says that an isolated mechanical system, starting in almost any initial state, will return arbitrarily close to that initial state at some time in the future. This means that, if a gas starts in an initial non-equilibrium state and evolves towards equilibrium in accordance with the H-theorem, then, by Poincaré's theorem, it must return to its initial non-equilibrium state at some later time. As the system evolves back towards the initial state, H must be increasing, in contradiction to the H-theorem.

Boltzmann's reply to these objections was based upon the assumptions made in the derivation of his equation and it helped to clarify the nature of those assumptions as well as to establish the range of validity of the equation. In deriving the Boltzmann equation, it was necessary to give an expression for the collisional effects in terms of the single particle distribution function. This "Stosszahlansatz" was statistical in nature. It essentially said that the number of binary collisions in a unit time may be found from the single particle distribution function by assuming that the particles' positions and momenta are uncorrelated. Given this lack of correlation as an initial condition, then, immediately after this instant, the Boltzmann equation holds and the H-theorem asserts that H is decreasing. It is only by assuming the Stosszahlansatz to hold at all times that H is found to be always decreasing. Actually, if the positions and momenta are uncorrelated before a binary collision, then that collision will introduce some correlation, and the Stosszahlansatz will no longer be exactly valid. Thus, the Stosszahlansatz is an approximation which will be valid only for a short period after the initial instant. This disposes of the Poincaré recurrence paradox since the recurrence time in macroscopic systems is very large while the H-theorem is valid only for small times. For example, for 10^{18} molecules in 1 cm^3 to return to their initial state with an accuracy of $|\Delta\vec{r}| \sim 10^{-7}$ cm and $|\Delta\vec{v}| \sim 10^{2}$ cm/sec takes a time of the order of $10^{10^{19}}$ years (cf. Wu [34]).

The effect of correlations, which are ignored in the H-theorem, is easily seen in the existence of fluctuations about equilibrium. If there are never any correlations, then the H-theorem is valid and $H(t)$ is constant so the system stays in equilibrium forever. What actually happens is that correlations in the system build up and result in a fluctuation, which is not described by the Boltzmann equation.

§10.2 A Simple Model and General *H*-theorems

In order to see how the "arrow of time", to which Loschmidt objected, arises, it is illustrative to study the exact evolution of a simple system. This model, which has recently been studied by several authors (cf. Blatt [35], Hobson [36], Lee [37], Kreuzer and Teshima [38]), shows clearly the role of statistical assumptions and initial conditions in irreversibility.

The model consists of a point particle which is trapped inside a one-dimensional box of length L, with perfectly elastic walls. If, at an initial time t=0 the system is described by a distribution function f(x,p,t=0), then the Liouville equation may be solved to find f(x,p,t). Actually doing so demonstrates that the evolution of f is irreversible in the sense that f will never return to its initial value. Also, it will be noted that the reduced distributions

$$h(p,t) = \int_0^L f(x,p,t)dx \qquad (10.12)$$

and

$$g(x,t) = \int_{-\infty}^{\infty} f(x,p,t)dp \qquad (10.13)$$

can monotonically relax to their final "equilibrium" values.

The Liouville equation is most easily solved if one takes care of the original boundary conditions by replacing the original distribution $f(x,p,o)$ by $\overline{f}(x,p,o)$ defined by

$$\overline{f}(x,p,o) = f(x,p,o) \quad \text{for} \quad 0 \leq x \leq L$$

and

$$\overline{f}(-x,-p,o) = \overline{f}(x,p,o) \quad ,$$

$$\overline{f}(x+2L,p,o) = \overline{f}(x,p,o) \quad . \tag{10.14}$$

The solution for the new function $\overline{f}$ is the same as that for f in the original problem (for $0 \leq x \leq L$) and, since the particles are free,

$$\overline{f}(x,p,t) = \overline{f}(x - \frac{pt}{m},p,o) \quad . \tag{10.15}$$

The solution may be expressed as a Fourier series in x. Before giving the solution, it is useful to specialize to an initial distribution in which the positions and momenta are uncorrelated. This condition is analagous to Boltzmann's Stosszahlansatz and may be expressed as

$$f(x,p,o) = g(x,o)h(x,o) \tag{10.16}$$

where g and h are normalized to one. With this initial condition, it is easily found that

$$f(x,p,t) = h(p,o)\{\frac{1}{L} + \sum_{n=1}^{\infty} a_n \cos[\frac{n\pi}{L}(x - \frac{pt}{m})]\} \tag{10.17}$$

where

$$a_n = \frac{2}{L}\int_0^L g(x,o)\cos(\frac{n\pi x}{L})dx \quad . \qquad (10.18)$$

To see the irreversible behavior of f(x,p,t) it is easiest to consider a specific case where

$$g(x,o) = \frac{1}{L}[1 + \cos(\frac{2\pi x}{L})] \quad . \qquad (10.19)$$

This is chosen to reduce the Fourier series to a single term. Also, let

$$h(p,o) = (\frac{\beta}{2\pi m})^{1/2} e^{-\beta\frac{p^2}{2m}} \qquad (10.20)$$

in order to make the evolution to a uniform final distribution monotonic. With these choices,

$$f(x,p,t) = \frac{1}{L}(\frac{\beta}{2\pi m})^{1/2} e^{-\beta\frac{p^2}{2m}}\{1 + \cos[\frac{2\pi}{L}(x - \frac{pt}{m})]\} \quad , \qquad (10.21)$$

$$h(p,t) = (\frac{\beta}{2\pi m})^{1/2} e^{-\beta\frac{p^2}{2m}} \quad , \qquad (10.22)$$

and

$$g(x,t) = \frac{1}{L}\{1 + e^{-\frac{2\pi^2 t^2}{mL^2\beta}} \cos(\frac{2\pi}{L} x)\} \quad . \qquad (10.23)$$

The momentum distribution h(p,t) remains Maxwellian while the position distribution g(x,t) relaxes monotonically to the final distribution

$$g(x,\infty) = \frac{1}{L} \quad . \qquad (10.24)$$

Although both $h(p,t)$ and $g(x,t)$ tend to specific final values as $t\rightarrow\infty$, it is not true that the system tends to equilibrium which would be given by

$$f^{eq}(x,p) = h(p,\infty)g(x,\infty) \quad . \qquad (10.25)$$

Instead, correlations between p and x have developed. This is seen in the term $\cos[\frac{2\pi}{L}(x-\frac{pt}{m})]$ which cannot be separated into a product of functions of x and p alone. With other choices of $g(x,o)$ and $h(p,o)$, there would be similar effects, although the decays to $g(x,\infty)$ and $h(p,\infty)$ would not, in general, be monotonic.

Thus, it is seen that specification of the initial condition as a continuous distribution is sufficient to yield irreversible behavior, even when the equation of motion is invariant under the substitutions $\vec{p} \rightarrow -\vec{p}$ and $\vec{t} \rightarrow -\vec{t}$. This is not a contradiction since, by specifying the initial condition at t=0 and considering the evolution for t>0, an "arrow of time" has been introduced. If the initial condition was stated as a condition on $f(x,p,t)$ at some time t_0 in the future, the system would irreversibly evolve toward that "initial" condition and away from equilibrium. The supposed reversibility paradox of Loschmidt is thus explained. If a statistical initial condition is introduced at t=0, in order for the distribution function to evolve so that it is the same at t as at -t, it is necessary that the initial condition as well as the equation of motion be time reversal invariant. If a system evolves irreversibly toward equilibrium and away from a given initial state and then, at some finite time, all velocities are reversed, certainly the system will evolve away from equilibrium back to the initial state. This is not a contradiction, but simply a reflection of the fact that the initial conditions in the original and in the time-reversed situations were different. In the above example, if the initial distribution was assumed to be uncorrelated, then the initial condition for the system in which the velocities were reversed at a later time is not the original uncorrelated distribution function, but rather the complex, correlated distribution to which the original system had evolved.

It should be noted that, for a continuous initial distribution, Zermelo's paradox does not occur since Poincaré's theorem on recurrence cycles applies only to isolated systems in a specific initial state. In the above example, this is clear since, for a single system in a particular state, the motion is periodic with period depending on the initial momentum. Thus, although the original system has a Poincaré recurrence cycle, the statistical distribution never returns to its initial value.

This is easily understood in terms of ensemble theory. If an ensemble of N identical systems is visualized, each member of which is in a possibly different, but still precise, initial state, then the corresponding distribution function is not continuous but rather consists of a series of N delta functions, one for each member of the ensemble. For finite N, the ensemble and therefore the distribution function has a Poincaré recurrence time since the ensemble may be regarded as an isolated supersystem obeying classical dynamics. The recurrence time, however, is much longer for the ensemble than for a single system and it increases rapidly with large N. In the limit $N \to \infty$, the recurrence time will tend to infinity and the distribution function will become continuous. Thus, the continuous initial distribution in the above example corresponds to the limit of an infinitely large ensemble which, of course, has no finite recurrence time.

Although this discussion clarifies the role of statistical distributions in irreversible behavior, it does not provide any justification for the initial introduction of such continuous distributions. Any justifications of ensemble theory based on ergodic theorems apply only to equilibrium and are therefore not useful in studying irreversible behavior.

It is often stated that a statistical distribution is a result of a lack of knowledge of the precise initial conditions. Since it is not practical to measure the position and momentum of each particle in a large system, in practice, an initial condition consists of specifying a few measureable parameters such as energy. All of the other parameters of the system are unspecified and this uncertainty means that a statistical initial condition is the best that can be achieved. This, however, has the perhaps unsatisfactory interpretation that the distribution function and thus also the physical averages which are calculated from it depend upon the precision with which the initial conditions are known. This point of view has a very precise and clear formulation in terms of information theory (cf. Katz [39]). The information theoretical approach also gives a clear interpretation of such procedures as coarse-graining.

In information theory, it is assumed that a certain amount of information is available from experimental measurements. That not everything is known about the system is reflected in the fact that the best possible description of the system is in terms of a distribution function ρ, whose evolution is governed by the Liouville equation,

$$\frac{\partial \rho}{\partial t} = \{H_N, \rho\} \quad . \tag{10.26}$$

The function I, defined by

$$I = -\mathcal{H}(t) = -\int \rho(\vec{x}_1,\dots,\vec{x}_N,t)\ln\rho(\vec{x}_1,\dots,\vec{x}_N,t)\,d^6\vec{x}_1\dots d^6\vec{x}_N \quad , \tag{10.27}$$

is a measure of the information missing. The function $\mathcal{H}(t)$ is known in statistical mechanics as the Gibbs fine-grained $\mathcal{H}$-quantity. Gibbs attempted to show that

$$\frac{d\mathcal{H}}{dt} \leqslant 0 \tag{10.28}$$

for an isolated system. However, using the Liouville equation, it can be shown that $\mathcal{H}(t)$ is a constant. This has the interpretation that, as long as the system is isolated and no new measurements are made, no information is gained or lost about the system.

In the above example of a particle in a one-dimensional box, the qualitative behavior of the information function is clear. Initially,

$$h(p,o) = \left(\frac{\beta}{2\pi m}\right)^{1/2} e^{-\beta\frac{p^2}{2m}} \tag{10.20}$$

and

$$g(x,o) = \frac{1}{L}\left[1 + \cos\left(\frac{2\pi x}{L}\right)\right] \tag{10.19}$$

and the information about momentum is a minimum, since (10.20) minimizes the momentum information function

$$\mathcal{H}_p(t) = \int_{-\infty}^{\infty} h(p,t)\ln h(p,t)\,dp \quad . \tag{19.29}$$

For a proof that $H_p(t)$ is minimized by h(p,o), see a discussion of Boltzmann's H-theorem (e.g. Huang [40]). However, the information about the positions is not a minimum since, for example, initially there can be no particles at x=L/2. As time progresses, $H_p(t)$ stays constant at its minimum value while the information about position tends monotonically to its minimum value (corresponding to $g(x,\infty)=\frac{1}{L}$). Since Liouville's theorem ensures that no information is lost, information must have been gained somewhere in the system. In fact, initially there was no information about correlations between x and p and, as time progresses, these correlations increase. Thus, all the lost information regarding position appears as information on correlations between x and p. With other initial conditions, similar behavior would be observed. Information would be transferred between position, momentum, and correlations, with a net irreversible tendency for g(x,t) and h(p,t) to go to their final minimum information values and for correlations to build up. That the final state still contains all the original information can be seen by performing the gedanken experiment of Loschmidt. Reversing all velocities will cause the system to return to its initial state with all information intact. This also implies that the system can never reach a true equilibrium since the correlations persist forever. The true equilibrium is given by

$$f^{eq}(x,p) = \frac{e^{-\beta H}}{\int_{-\infty}^{\infty} e^{-\beta H}dxdp}$$

$$= \frac{1}{L}\left(\frac{\beta}{2\pi m}\right)^{1/2} e^{-\beta\frac{p^2}{2m}} \quad \text{for} \quad 0 \leq x \leq L$$

$$= 0 \qquad \text{otherwise}$$

(10.30)

which is just $f^{eq} = g(x,\infty)h(p,\infty)$ (10.31)

without any correlations.

If an observer is limited to certain types of measurements (e.g. he can measure h(p,t) and g(x,t) but not correlations), then the above system's behavior will appear to him to be an irreversible approach to equilibrium. With these restrictions on possible observations, the quasi-equilibrium, which includes all correlations, is indistinguishable from the true equilibrium. Such an approach to the problem of irreversibility is often justified by a procedure known as coarse-graining. It is argued that all physical measurements are actually averages over some finite time τ. Thus, f(x,p,t) is not accessible to direct measurement but only such quantities as

$$\overline{f}(x,p,t) = \frac{1}{\tau}\int_{t}^{t+\tau} f(x,p,t')dt' \quad . \qquad (10.32)$$

In the above example, if τ is chosen to be one period of oscillation of the cosine factor in eqn. (10.21)

$$\tau = \frac{mL}{2\pi p}\,\frac{1}{t} \quad , \qquad (10.33)$$

then

$$\overline{f}(x,p,t) = f^{eq} \quad . \qquad (10.34)$$

Note also that the coarse-graining time τ is inversely proportional to t so that any finite size for the coarse graining time will give the equilibrium distribution at some time $t \alpha \frac{1}{\tau}$. Thus, since the correlations oscillate with a shorter and shorter period as $t \to \infty$, measurements with an arbitrary but finite precision will "smear out" the correlations eventually.

The statement that no measurement of finite precision can distinguish between true and quasi-equilibrium is often cited as a complete explanation of irreversible behavior. It is claimed that the information destroyed by the formal procedure of coarse-graining is truly lost since it is impossible to perform a measurement to retrieve it.

Historically, coarse-graining was introduced into statistical mechanics by the Ehrenfests to derive an H-theorem for the Gibbs fine-grained H-quantity defined by eqn. (10.27). Using $\vec{X}$ to denote all 6N coordinates, the H-quantity is

$$H(t) = \int_{\Gamma} \rho(\vec{X}) \ln \rho(\vec{X}) d^{6N}\vec{X} \tag{10.35}$$

where the integral is over all Γ-space. Then, it can be shown that

$$\frac{dH}{dt} = 0 \quad . \tag{10.36}$$

However, by coarse-graining (i.e. averaging) over cells in Γ-space of size ΔΓ, a new quantity $\overline{\rho}(\vec{X},t)$ may be defined as

$$\overline{\rho}(\vec{X},t) = (\Delta\Gamma)^{-1} \int_{\Delta\Gamma} \rho(\vec{X},t) d^{6N}\vec{X} \tag{10.37}$$

where ΔΓ includes the point $\vec{X}$. This new function does yield an H-theorem of sorts.

Defining a coarse-grained H-quantity by

$$\overline{H}(t) = \int_{\Gamma} \rho(\vec{X},t) \ln \overline{\rho}(\vec{X},t) d^{6N}\vec{X} \tag{10.38}$$

gives

$$H(t) \geqslant \overline{H}(t) \tag{10.39}$$

for all times. This can be proven using the mathematical result that, given any two functions $\rho_1(\vec{X},t)$ and $\rho_2(\vec{X},t)$ normalized such that

$$\int_{\Gamma} \rho_1(\vec{X},t) d^{6N}\vec{X} = \int_{\Gamma} \rho_2(\vec{X},t) d^{6N}\vec{X} \quad , \tag{10.40}$$

then,

$$\int_{\Gamma} \rho_1 \ln \rho_1 d^{6N}\vec{X} \geqslant \int_{\Gamma} \rho_1 \ln \rho_2 d^{6N}\vec{X} \tag{10.41}$$

where the equality holds only for $\rho_1 = \rho_2$ (for a proof, see Isihara [25]). Upon noting that

$$\int_{\Gamma} \bar{\rho}\, d^{6N}\vec{X} = \int_{\Gamma} \rho\, d^{6N}\vec{X} \quad , \tag{10.42}$$

the result (10.39) follows immediately.

If then an initial fine-grained distribution ρ is chosen constant over each of the small cells, then

$$\bar{H}(o) = H(o) \quad . \tag{10.43}$$

Noting that $H(t) = H(o) = \bar{H}(o)$ gives

$$\bar{H}(o) \geq \bar{H}(t) \tag{10.44}$$

which is a generalized H-theorem. Gibbs argued that an initial distribution would generally spread out in a highly complex manner throughout the entire Γ-space, forming long "filaments" of density ρ. He believed that eventually these filaments would fill the entire available region of Γ-space with a constant average density $\bar{\rho}$, corresponding to equilibrium. However, this is not proven and it is not even true generally that

$$\frac{d\bar{H}}{dt} \leq 0 \quad . \tag{10.45}$$

Such a method of deriving an H-theorem seems to be a mathematical construction and requires some physical justification. The usual justification is that only coarse-grained quantities of some sort are measured. This is a valid explanation of irreversibility only provided that it is impossible to extract or use all of the fine-grained information. Although this may seem reasonable due to the highly complex nature of the correlations, Blatt [35] has argued that it is not necessary to actually measure the fine-grained information to demonstrate its existence. He gives an illustrative example of a spin system which is analagous to the time-reversal gedanken experiment of Loschmidt (cf. Hahn [41]).

Blatt's system, which occurs in nuclear magnetic resonance studies, consists of a large number of magnetic dipoles, associated with spins, in a uniform external magnetic field $\vec{H}_0$. The magnetic moments are normally aligned parallel to $\vec{H}_0$. At $t=0$, an oscillating magnetic field $\vec{H}_1(t)$ is applied to the sample at right angles to $\vec{H}_0$ and at the Larmor precession frequency. This perturbation is applied quickly enough and for just long enough that its net effect is to align all the moments in the same direction in the plane perpendicular to $\vec{H}_0$. When the perturbing field is removed, the moments all precess in the external field at their respective Larmor frequencies. Due to small inhomogeneities in $\vec{H}_0$, these frequencies are slightly different and, as a result, the various moments rapidly get out of phase. This decay of alignment after the removal of $\vec{H}_1(t)$ is observed by measuring the magnetization of the sample. When the spins are aligned, there is a net magnetization which precesses at the Larmor frequency and which is observed to decay rapidly, in a time t_1, as the spins become unaligned. Thus, the spin system appears to decay irreversibly to equilibrium in a time t_1. If, however, at a time $t_2 >> t_1$, another pulse of the same magnitude and direction as $\vec{H}_1(t)$ is applied for twice the duration of $\vec{H}_1(t)$, the net effect is to reflect all the spins about a plane parallel to $\vec{H}_0$. This is directly analogous to reversing all the velocities in a system of particles. The correlations between the slightly differing precession frequencies and the precession angles then result in the spins realigning themselves at a time $2t_2$ and giving a measurable peak in the magnetization. If the perturbing pulse is reapplied at intervals $2t_2$ apart, the spins repeatedly realign to give more peaks in magnetization. These peaks, however, decrease in magnitude since some of the spins interact thermally with the lattice, which acts as a thermal reservoir, and cease to precess in the plane of $\vec{H}_1(t)$. Eventually, in a time τ, this external interaction destroys all correlations and the system does not show any more magnetization peaks. Thus, the true irreversible behavior occurs over a time τ governed by the interaction of the system with its surroundings. The system can reach a quasi-equilibrium in a time t_1 much shorter than τ. This t_1 is a time in which the correlations become so complex that the usual macroscopic measurements cannot detect them. These correlations are still present and can be determined from fine-grained measurements or may be seen in the evolution of the time-reversed state away from apparent equilibrium.

These conclusions may be carried over to general thermodynamical systems. Starting from a non-equilibrium initial distribution, a system will rapidly evolve, via its own dynamics, to a quasi-equilibrium state in which the measured macroscopic parameters of the system have their equilibrium values but in which there are some quantities which are far from equilibrium. In practice, these far-from-equilibrium quantities will involve complex correlations and will seldom, if ever, be measured. Thus, the distribution function of an isolated system will show irreversible evolution to an apparent equilibrium but, due to the inability of an isolated system to "forget", it is only through external interactions that a true equilibrium may be attained. This process will generally take longer than the time period for the attainment of the above quasi-equilibrium.

Although these ideas form a more or less consistent interpretation of irreversibility via statistical distribution functions, the comments of Chapter VIII regarding the manner in which statistical considerations are introduced are still valid. Although the introduction of continuous distribution functions may be accepted as a fundamental postulate of non-equilibrium statistical mechanics, this is somewhat unsatisfactory in view of the lack of any non-equilibrium analogue of the ergodic theorem to connect the distribution function (or ensemble) averages to the "real life" observations on single systems.

As discussed in Chapter VIII, the introduction of statistical ideas may be regarded as necessary in order to take into account the interaction of a system with its surroundings. It is consistent with the usual ideas of a thermal reservoir to assume that interaction with the system does not change the properties of the reservoir in any measurable way. Thus, information may be transferred to the reservoir, causing a net loss of information from the system. This allows the system to evolve to a true equilibrium in a manner that is not possible for isolated systems. This may be regarded as a random thermal interaction which acts to destroy the fine-grained correlations present in an isolated system. It should be remembered that this interaction is truly random only for times much less than the Poincaré recurrence time of the composite system; however, this time may be made as long as desired by considering a large thermal reservoir.

This discussion agrees with Blatt's analysis of spin-echo experiments. Blatt analyses the particle in a box model with a thermal interaction, represented by a random "thermal" velocity of the walls. This model has since been studied in more detail by Kreuzer and Teshima

[38]. Blatt concludes that the system has two time scales associated with the thermal interaction. In the shorter time τ_1, the random influence of the walls has destroyed the detailed information in the system and, in the larger time τ_2, the thermal interaction has exchanged enough energy between the particles and the walls to bring the system into thermal equilibrium with the walls. This time τ_2 corresponds to the third time scale of Bogolyubov in which the system goes to complete equilibrium. However, Blatt's time τ_1 does not correspond to Bogolyubov's time scale for the decay of correlations since there are no interparticle interactions to cause such a decay. Neither is there a kinetic stage in this model because of this lack of interactions. Instead, the separation between Blatt's two time scales depends upon the dynamics of the wall-particle collisions (specifically upon the ratio of the mass of a particle to the mass of a constituent of the wall) and is not fundamental as are the time scales of Bogolyubov.

It would seem that there are some essential features missing in Blatt's model. Some of these are due to the fact that the model contains no interactions so a particle must collide with the walls several times in order to have its velocity randomized. This leads to a time scale for relaxations which is much longer than expected. Although explicitly adding such an interparticle interaction would allow sharing of thermal energy and hence a randomization of the system in a shorter time, of order of the mean time between collisions, this would render explicit solutions impossible. The lack of interparticle collisions also leads to another unphysical feature of the model, first noticed by Kreuzer and Teshima [38]. That is, as time progresses, the particles spend most of their time in states of very low velocity resulting from collisions with the wall which nearly stop the particles. Thus, following the evolution of the system for large times rather than for large numbers of collisions does not yield an evolution to equilibrium between the particles and the walls. This feature could be remedied either by an artificial restriction which would exclude such collisions or by a more detailed analysis of the dynamics of wall-particle collisions. However, such a model would still not show the behavior predicted by Bogolyubov's analysis.

It is obvious that such a model (corresponding to an ideal gas) cannot show distinguishable time scales for evolution but will simply show a complete randomization of each particle's motion in a single time τ_2.

However, even though interactions are necessary to have evolution and randomization in the time scales introduced by Bogolyubov, the thermal interaction is seen to provide a suitable mechanism to convert a quasi-equilibrium state into a true equilibrium one and it is plausible that the effect of interparticle collisions would be to greatly accelerate this randomization.

§10.3 Quantum Systems and Master Equations

In general, the introduction of quantum mechanics does not significantly change the problem of irreversibility. Most of the difficulties of the classical theory have direct analogues in quantum mechanics. The quantum version of the Liouville equation (von Neumann equation) is time-reversal invariant and has an interpretation in terms of ensembles similar to that of its classical counterpart.

The von Neumann equation may be written as

$$\frac{\partial\rho}{\partial t} = \frac{1}{i\hbar}[H,\rho] \qquad (10.46)$$

and may be found either from the correspondence $\{,\}\rightarrow\frac{1}{i\hbar}[,]$ between classical and quantum mechanics or directly from the definition of ρ in Chapter VIII using the Schrödinger equation.

In quantum statistics, there have been several attempts to derive a "master equation" to describe the time evolution of statistical quantities. If a system is described by a Hamiltonian $H = H_0 + H_i$ where H_0 describes the isolated system and H_i describes an interaction which causes transitions between eigenstates of H_0, then, in the notation of Chapter VIII,

$$\psi^k(\vec{x}_1,\ldots,\vec{x}_N,t_0) = \sum_n c_n^k(t_0)\phi_n(\vec{x}_1,\ldots,\vec{x}_N) \qquad (10.47)$$

where k denotes the k^{th} member of an ensemble and the ϕ_n are a complete set of eigenstates of H_o. Then, from quantum mechanics, it is known that

$$\psi^k(t) = \sum_{n,m} U_{nm}(t,t_o)c_m^k(t_o)\phi_n$$

$$= \sum_n c_n^k(t)\phi_n \qquad (10.48)$$

where U is the unitary time development operator. Thus,

$$c_n^k(t) = \sum_m U_{nm}(t,t_o)c_m^k(t_o) \qquad (10.49)$$

and

$$c_m^{k*}(t)c_n^k(t) = \sum_{i,j} U_{ni}(t,t_o)U_{mj}^*(t,t_o)c_i^k(t_o)c_j^{k*}(t_o) \quad . \qquad (10.50)$$

Upon averaging over the N members of an ensemble,

$$\rho_{nn}(t) = \frac{1}{N}\sum_{k=1}^{N} c_n^{k*}(t)c_n^k(t)$$

$$= \sum_{i,j} U_{ni}U_{nj}^* \frac{1}{N}\sum_{k=1}^{N} c_i^k(t_o)c_j^{k*}(t_o) \quad . \qquad (10.51)$$

At this point, the assumption of random phases is made. This assumption, justified by the impossibility of measuring the phases associated with the wave functions of the members of the ensemble, asserts that

$$\rho_{ij}(t_o) \equiv \frac{1}{N}\sum_{k=1}^{N} c_i^k(t_o)c_j^{k*}(t_o)$$

$$= 0 \quad \text{for } i \neq j \quad . \qquad (10.52)$$

This assumption is an initial condition for ρ_{ij} and plays a role similar to the assumption that the fine-grained distribution is initially uniformly spread over several cells in Γ-space in the discussion of classical coarse-graining. Under this assumption,

$$\rho_{nn}(t) = \sum_i |U_{ni}(t,t_o)|^2 \rho_{ii}(t_o) \quad . \qquad (10.53)$$

Writing

$$|U_{ni}(t,t_o)|^2 = A_{ni}(t_o,\Delta t) \qquad (10.54)$$

where $\Delta t = t-t_o$ and A_{ni} is then the probability of a transition from a state i to a state n in the interval Δt. Letting $P_i(t)$ denote the probability that a system is in state i at time t gives

$$P_n(t_o+\Delta t) = \sum_i A_{ni}(t_o,\Delta t)\, P_i(t_o) \quad . \qquad (10.55)$$

That this expression contains a built in irreversibility may be shown as follows. From the interpretation of the A_{ni} as transition probabilities, it is obvious that

$$\sum_n A_{ni} = \sum_i A_{ni} = 1 \qquad (10.56)$$

and that

$$0 \leqslant A_{ni} \leqslant 1 \quad . \qquad (10.57)$$

If the inverse of the matrix A_{ni} exists, then

$$P_i(t_o) = \sum_n A^{-1}_{in}\, P_n(t_o+\Delta t) \quad . \qquad (10.58)$$

However, since $\sum_n A^{-1}_{in} A_{nj} = \delta_{ij}$, it follows that not all of the elements A^{-1}_{in} can be between zero and one unless $A_{nj} = \delta_{nj}$. Thus, if $A_{nj} \neq \delta_{nj}$, the elements of A^{-1} cannot be interpreted as transition probabilities. The origin of this irreversibility is the assumption that, at t_o, the density matrix is diagonal. Then, unless $|U_{ni}|^2 = \delta_{ni}$, for $t>t_o$ there will be some off-diagonal elements in $\rho_{ij}(t)$. Thus, for $t>t_o$, no simple equation involving only the transition probabilities between states can be valid.

Starting from a diagonal $\rho_{ij}(t_o)$, it is possible to prove a type of *H*-theorem, known as Klein's lemma. To do so, define

$$H(t) = \sum_i P_i(t)\ln P_i(t) = \sum_i \rho_{ii}(t)\ln\rho_{ii}(t) \qquad (10.59)$$

and

$$Q_{ji}(t_o,\Delta t) = \rho_{jj}(t_o)[\ln\rho_{jj}(t_o) - \ln\rho_{ii}(t_o+\Delta t)-1] + \rho_{ii}(t_o+\Delta t) \quad . \qquad (10.60)$$

Due to the inequality

$$x(\ln x - \ln y) - x+y \geqslant 0 \qquad (10.61)$$

with the equality valid only for x=y,

$$Q_{ij}(t_o,\Delta t) \geqslant 0 \quad . \qquad (10.62)$$

Therefore,

$$\sum_{i,j} A_{ij}(t_o,\Delta t)Q_{ij}(t_o,\Delta t) = \sum_j \rho_{jj}(t_o)\ln\rho_{jj}(t_o)$$

$$- \sum_i \rho_{ii}(t_o+\Delta t)\ln\rho_{ii}(t_o+\Delta t) \geq 0 \qquad (10.63)$$

since $A_{ij} \geq 0$ and where the fact that

$$\sum_j \rho_{jj}(t_o) = \sum_i \rho_{ii}(t_o+\Delta t) \qquad (10.64)$$

was used.

Thus, using the definition (10.59) of $H(t)$ gives

$$H(t_o+\Delta t) \leq H(t_o) \quad . \qquad (10.65)$$

This H-theorem results from the spreading of the density matrix away from the diagonal form which it had at t_o. Even with this initial condition, this does not imply a monotonic decay of $H(t)$, since, for $t>t_o$, $\rho_{ij}(t)$ is no longer diagonal and eqn. (10.59) no longer holds. The "master equation" as derived by Pauli, assumes that (10.59) is valid at all times. On this basis, it can be shown that $H(t)$ decays monotonically to an equilibrium value. However, this ignores the non-diagonal part of ρ_{ij} and may be regarded as analogous to the Stosszahlansatz used in deriving classical kinetic equations.

The relationship (10.65) is more fundamental since it assumes diagonality only at an initial instant. Although the weak nature of this H-theorem and its derivation from an initial condition on ρ_{ij} are suggestive of the classical coarse-grained *H*-theorem, no coarse-graining procedure was introduced in deriving (10.65). This *H*-theorem has no classical analogue. This is a reflection of the fact that $\sum_i \rho_{ii} \ln \rho_{ii}$ is not the trace of any operator and hence is basis dependent. Thus, except at the instant when ρ_{ij} is diagonal, it does not represent an observable. Therefore, this *H*-theorem does not provide an explanation of observed macroscopic irreversibility.

In order to remove this objection, it is necessary to find the conditions under which ρ_{ij} may be regarded as diagonal. By considering the distinction between macroscopic and microscopic operators, it is possible to perform a quantum coarse-graining procedure and, from this, to derive a valid *H*-theorem (cf. Jancel [42]).

Consider two quantum observables A and B. If these do not commute,

$$\Delta A \Delta B \geqslant \frac{1}{2} \left| \overline{[A,B]} \right| \qquad (10.66)$$

where the bar denotes the quantum mechanical average. However, all the macroscopic observables can be measured simultaneously, so it is necessary to construct commuting macroscopic observables corresponding to the microscopic observables A and B.

To see how to do this, consider the energy of a system. A macroscopic measurement of E will, in general, have an experimental uncertainty δE. For a large system, δE will be much larger than the separation between the eigenvalues E_i of the Hamiltonian, H. Thus, a macroscopic measurement of E will not distinguish between eigenvalues of H, but only between groups of eigenvalues of nearly the same energy. Thus, one is led to introduce cells $e^{(\alpha)}$ of size δE and a macroscopic energy operator H_M, defined by

$$H_M = \sum_\alpha \varepsilon_\alpha \, P^{(\alpha)} \qquad (10.67)$$

where ε_α is the average energy in the cell $e^{(\alpha)}$ and where $P^{(\alpha)}$ is the projection operator for cell $e^{(\alpha)}$, defined by

$$P^{(\alpha)} = \sum_{i=1}^{S(\alpha)} P_i \qquad (10.68)$$

where P_i is the projection operator for the eigenvalue E_i in the cell $e^{(\alpha)}$ and where $S(\alpha)$ is the number of eigenvalues in that cell. Thus, the macroscopic energy operator has a greater degeneracy for each macroscopic energy eigenvalue ε_α than does the operator H.

In the energy representation, both H and H_M are diagonal, with eigenvalues E_i and ε_α respectively. Consider now another microscopic operator A which does not commute with H. Then, in the same representation,

$$[A,H]_{ij} = (E_i - E_j)A_{ij} \sim \Delta E \Delta A \quad . \qquad (10.69)$$

The macroscopic measurements satisfy $\delta E >> \Delta E$ and $\delta A >> \Delta A$ so

$$\delta E \delta A >> \Delta E \Delta A \quad . \tag{10.70}$$

Taking $E_i - E_j > \delta E$ in eqn. (10.69) gives

$$\delta E A_{ij} << \delta E \delta A \tag{10.71}$$

or
$$A_{ij} << \delta A \tag{10.72}$$

for
$$E_i - E_j > \delta E \quad . \tag{10.73}$$

Thus, matrix elements A_{ij} corresponding to energy differences greater than δE can be neglected. This means that all non-negligible elements of A_{ij} are grouped in a band along the diagonal which is narrow compared to the size of the cells $e^{(\alpha)}$.

Dropping negligible elements, a macroscopic operator A_M is constructed which consists of a set of submatrices each corresponding to a cell $e^{(\alpha)}$. This operator commutes with H_M and so the two may be simultaneously diagonalized. Continuing in this manner for other operators B, C, etc., it is possible to construct a set of commuting macroscopic observables which may all be diagonalized in a representation in which H_M is diagonal. This is possible because of the coarse-grained nature of macroscopic observations.

In order to derive an *H*-theorem, define a coarse-grained probability density $\rho_{M_{ij}}^{(\alpha)}$ by averaging ρ_{ii}

over the cells $e^{(\alpha)}$. That is

$$\rho^{(\alpha)}_{M_{ij}} = \delta_{ij} \frac{1}{S(\alpha)} \sum_{i=1}^{S(\alpha)} \rho^{(\alpha)}_{ii} \qquad (10.74)$$

where $\frac{1}{S(\alpha)} \sum_{i=1}^{S(\alpha)} \rho^{(\alpha)}_{ii}$ is the statistical weight of cell $e^{(\alpha)}$. Define an H-quantity by

$$\begin{aligned} \overline{H}^{(\alpha)}(t) &= \mathrm{Tr}(\rho^{(\alpha)}_{M} \ln \rho^{(\alpha)}_{M}) \\ &= \sum_{i=1}^{S(\alpha)} \rho^{(\alpha)}_{M_{ii}} \ln \rho^{(\alpha)}_{M_{ii}} \quad . \end{aligned} \qquad (10.75)$$

Then, in a manner similar to the derivation of eqn. (10.65), it may be shown that

$$\overline{H}^{(\alpha)}(t) \leq \overline{H}^{(\alpha)}(o) \qquad (10.76)$$

where it has been assumed that at t=0, the fine-grained quantity $\rho^{(\alpha)}_{ij}$ is equal to the coarse-grained quantity $\rho^{(\alpha)}_{M_{ij}}$. This coarse-grained $\overline{H}$-theorem now has some physical meaning since, by defining appropriate macroscopic operators, it was possible to have $\rho^{(\alpha)}_{M}$ diagonal in a basis in which all the other macroscopic observables were diagonal and hence to prove an H-theorem for a quantity $\mathrm{Tr}(\rho^{(\alpha)}_{M} \ln \rho^{(\alpha)}_{M})$ which can represent an observable quantity since it is basis independent.

Thus, it seems that there are no fundamental differences introduced into discussions of irreversibility by quantum mechanics. Although the details may be different, the same difficulties occur in both the classical and quantum cases and the interpretations of results are similar.

In conclusion, it has been shown that if statistical considerations are introduced into the initial conditions of an isolated system, the Liouville equation describes irreversible evolution of the distribution function. However, the distribution function of an isolated system can never reach a true equilibrium, since all of the information about correlations is retained. If observations are restricted to ones that do not involve all the correlations, the system may reach a quasi-equilibrium. Even though some of the measured quantities in such an evolution may evolve monotonically to their equilibrium values, it will be possible at all times to find some usually complex quantities which are far from their equilibrium values. The fact that this is only a statistical irreversibility is obvious from the fact that a single isolated system must exhibit both fluctuations and a Poincaré cycle, while a continuous probability distribution may have neither. Unfortunately, there is no proof of the equivalence of ensemble averages and time averages for an isolated non-equilibrium system, so H-theorems which are based on continuous distributions, when applied to single systems, can yield only weak statements. Such an H-theorem might say that a single system will probably evolve to a state in which the measured, macroscopic parameters will approach their equilibrium values. Just how probable this is and the precise time scales and nature of the approach to equilibrium depends upon the dynamics of the system. This, however, is outside the range of purely statistical considerations. It seems that, although for short times and limited observations, the distinction between equilibrium and quasi-equilibrium is unimportant, the most satisfactory way in which to allow the establishment of a true equilibrium is via an external interaction.

CHAPTER XI

RESPONSE OF QUANTUM SYSTEMS TO MECHANICAL PERTURBATIONS

§11.1 Response Theory Formalism

In quantum statistical mechanics, there is one class of non-equilibrium processes which are of great interest and for which formal solutions are easily obtained. These are the processes in which the deviation from equilibrium is due to an external perturbation which can be described by a term in the Hamiltonian. This includes a wide variety of important physical processes. For example, the electrical currents in a system acted on by an external electrical field or the magnetization in an external magnetic field are commonly treated in this way. However, more general external perturbations, such as thermal interactions, cannot be described by a Hamiltonian and so cannot be treated. It is also obvious that this formalism does not describe or explain the irreversible evolution towards equilibrium discussed in the previous chapter, since the system is maintained in a non-equilibrium state by the application of an external perturbation.

Although the formalism of this chapter is usually applied to transport processes and then only in the lowest order (linear response), actually quite a large range of situations may be treated. In particular, the systems which can be described in this way have a Hamiltonian of the form

$$H = H_o + H'_t \qquad (11.1)$$

where H_o is the total Hamiltonian of the isolated system and H'_t is the time-dependent Hamiltonian of the interaction between the external perturbation and the original system. The theory of linear response was derived and applied by several authors (cf. Kubo [43], Nakano [44]) and the resulting formulae are commonly referred to as Kubo formulae.

Following Kubo, the perturbing Hamiltonian will be written as

$$H'_t = -\mu\ F(t)A \qquad (11.2)$$

where A is a Schrödinger operator, μ is a coupling constant, and F(t) is a c-number function of time. The perturbation is assumed to be turned on at some finite time so

$$F(t)\Big|_{t=-\infty} = 0 \quad . \qquad (11.3)$$

Starting from some initial distribution $\rho_o \equiv \rho(t=-\infty)$, the evolution of ρ is governed by the equation

$$i\hbar\ \frac{\partial\rho}{\partial t} = [H,\rho] = [H_o,\rho] - \mu F(t)[A,\rho] \quad . \qquad (11.4)$$

Since it is only the response of the system to the perturbation and not the detailed dynamics which is of interest, it will be assumed that the system is initially in a time-independent state. This means that $[\rho_0, H_0] = 0$ and ρ_0 will usually be chosen as a canonical distribution.

To simplify eqn. (11.4), convert all operators to the interaction picture. This is done by defining an interaction picture operator $B_I(t)$ corresponding to an arbitrary Schrödinger operator B by

$$B_I(t) = e^{i\frac{H_0 t}{\hbar}} B \, e^{-i\frac{H_0 t}{\hbar}} \quad . \tag{11.5}$$

Eqn. (11.4) then becomes

$$i\hbar \frac{\partial \rho_I(t)}{\partial t} = -\mu F(t)[A_I(t), \rho_I(t)] \tag{11.6}$$

which can be written as an integral equation,

$$\rho_I(t) = \rho_I(t=-\infty) + \left(-\frac{\mu}{i\hbar}\right) \int_{-\infty}^{t} dt_1 F(t_1)[A_I(t_1), \rho_I(t_1)] \tag{11.7}$$

and, since $[\rho_0, H_0] = 0$,

$$\rho_I(t=-\infty) = \rho_0 \quad . \tag{11.8}$$

Thus, iterating eqn. (11.7) gives (after transforming back from the interaction picture)

$$\rho(t) = \rho_0 + \sum_{n=1}^{\infty} \left(-\frac{\mu}{i\hbar}\right)^n \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \ldots \int_{-\infty}^{t_{n-1}} dt_n \{ e^{-i\frac{H_0 t}{\hbar}} F(t_1) \ldots F(t_n)$$

$$\times [A_I(t_1), [A_I(t_2), \ldots, [A_I(t_n), \rho_0] \ldots] e^{i\frac{H_0 t}{\hbar}} \} \quad . \tag{11.9}$$

This is the quantum analogue of eqn. (9.63) and is an exact formal solution. However, for increasing times, more and more terms must be kept to have an accurate expression for $\rho(t)$. Its practical value is also limited by the necessity of exactly calculating terms of the form

$$\{e^{-i\frac{H_o t}{\hbar}}[A_I(t_1),[A_I(t_2),\ldots,[A_I(t_n),\rho_o]\ldots]e^{i\frac{H_o t}{\hbar}}\}. \quad (11.10)$$

Even in the lowest order approximation,

$$\rho(t) = \rho_o - \frac{\mu}{ih}\int_o^t dt_1\{e^{-i\frac{H_o t}{\hbar}} F(t_1)[A_I(t_1),\rho_o]e^{i\frac{H_o t}{\hbar}}\}, \quad (11.11)$$

getting an explicit solution would be at least as complicated as solving for the entire evolution of the unperturbed system, since H_o describes the entire dynamics of the unperturbed system.

However, it is possible to derive important and quite general relations for the response of a system from this formalism. Examples of such relations, obtainable from linear response, are Onsager relations and fluctuation-dissipation relations. Before deriving these, some discussion of the validity and applicability of the formalism and of the conditions under which the series (11.9) may legitimately truncated, as in eqn. (11.11), is in order.

First, however, note that the above non-linear response theory is exactly equivalent to the perturbation theory familiar from many-body or quantum field theory. In those theories, a perturbation expansion is found for the unitary time development operator $U(t,t_o)$. This expansion is usually written as

$$U(t,t_o)=e^{-i\frac{H_ot}{\hbar}}\sum_{n=0}^{\infty}\left(-\frac{\mu}{i\hbar}\right)^n\int_{t_o}^{t}dt_1\int_{t_o}^{t_1}dt_2\ldots\int_{t_o}^{t_{n-1}}dt_n H_I'(t_1)\ldots H_I'(t_n)e^{i\frac{H_ot}{\hbar}} \quad . \qquad (11.12)$$

In many-body theory, the usual interpretation is that H_o describes a system of free particles and the perturbation H' is the interaction between particles. There is nothing, however, in the derivation of eqn. (11.12) which requires this and, if H_o is the total Hamiltonian of the unperturbed system and $H'$is the Hamiltonian of the external perturbation, then it can be shown that

$$\rho(t) = U(t,-\infty)\rho_o\, U^{\dagger}(t,-\infty) \qquad (11.13)$$

is exactly the same as the response theory expression (11.9) for $\rho(t)$ (cf. Bernard and Callen [45]). This equivalence is mentioned merely to emphasize that the linear and non-linear response theory is a purely dynamical perturbation theory even though it is applied to a statistical operator.

§11.2 Validity of Response Theory

From the derivation of the expansion (11.9) for $\rho(t)$, it is clear that the theory describes a system which is initially, at some time in the distant past, in equilibrium. This may be achieved, for example, by putting the system in contact with a thermal reservoir. Then, at some past time, the system is isolated from the reservoir, a perturbation H_t' is turned on, and the dynamical evolution of $\rho(t)$ is given by (11.9). Thus, in analysing the validity of linear or of non-linear response theory, there are two important questions. First, under what conditions is the isolated system of the perturbation theory a good approximation to the physically realizable situation? Secondly, for what conditions is the linear response expression (11.11) a sufficient approximation to the exact expression (11.9)?

In answer to the first question, it would seem that, in many physical systems, the thermal and other non-Hamiltonian interactions are vital. For example, in order to calculate the electrical conductivity of a metal at temperature T, it might seem reasonable to apply the preceding formalism. However, if the system is considered long after the applied field is shut off, there are two different situations. In the physically realizable system, the metal has returned to equilibrium at the same temperature T due to thermal interactions. However, in the isolated system to which the above response theory applies, the system will, for example, have absorbed energy in the form of Joule heating and so will not return to temperature T without the thermal interaction. Note that this is a macroscopic effect and that the microscopic motion will be drastically affected by even a very small thermal interaction.

Also, in such a conductivity experiment, another fundamental thermal interaction is present since, in order to maintain a steady current through the system, electrons are supplied from a battery at one end and are removed at the other. In order that this not significantly affect the results, it is necessary that the electrons in the system always have essentially the same distribution as those externally supplied. This further restricts the validity of response theory.

It is very important to note that it is only in the linear response term that the effects of thermal and of mechanical perturbations are independent. Thus, even if the thermal perturbations are such that they average out to give an insignificant contribution to the linear response, they may still interfere with the mechanical non-linear response.

This analysis is rather pessimistic and it is often possible to replace the physical system with an idealized isolated system to which response theory applies and which would be expected to show the same general features as the physical system. With such a system, the linear mechanical response may be calculated and, from experience, it is expected that the effect of small thermal fluctuations upon the conductivities would be small. For example, for electrical conductivity, the need for exchanging thermal electrons can be eliminated by considering a toroidal conductor and by adiabatically supplying a tangential electric field by changing the magnetic flux through the center of the toroid. In such a situation, it is reasonable to assume that the resultant conductivities are very nearly the same as for the usual conductivity experiments.

Although it could be argued that the isolated system described above could never reach a steady state due to the previously mentioned Joule heating effect, it can be shown that the first term in the Joule heating is quadratic in the applied field and hence, for small fields, it is expected that the system should reach a steady state long before Joule heating becomes important.

Thus, although it has not been proven, it is reasonable to expect that the linear response expressions are valid for systems which can be replaced by idealized isolated systems. This is not true of the higher order responses and, even for linear response, is only true to the extent that the macroscopic response is insensitive to thermal perturbations. Note that certain attempts to explicitly take into account the thermal perturbations, for example by introducing a time-dependent temperature, are not sufficient since they only account for a single macroscopic effect and not for the complex microscopic effects.

Concerning the question as to the range in which the linear approximation is sufficient, van Kampen [46] has given some arguments and estimates. Essentially, he claims that linear response is valid only for applied fields so small as to be physically unrealistic and that the observed linearity of responses, such as electrical currents, is the result of a complex average behavior and not of the linearization of the equation of motion for $\rho(t)$.

To justify this, he uses an example of a classical system of electrons which are scattered by a single impurity. If, as van Kampen does, one chooses a diameter of 10^{-8} meters for the impurity and an effective mass of 10^{-28} kilograms for the electrons, and agrees with his argument that, for linear response theory to be valid, the displacement of each electron's position from the position it would have in the absence of a perturbation must be much less than the diameter of the impurity, then

$$t^2 E_0 \ll 10^{-16} \text{ sec}^2 \text{ Volt/meter} \qquad (11.14)$$

for a static electric field E_0 and for a time t after the perturbing field was turned on. This criterion would be even more severe in a physical system where there are many impurities and where multiple scattering is possible. It must be realized that this criterion gives the range for which all phenomena, including transient responses, are linear in the applied field at time t. This ensures, for example, that if an isolated conductor is placed between the plates of a capacitor and, at t_0, a constant electric field is turned on, then, for the time given by criterion (11.14), all the motions are linear in the field strength E_0. This time is not a macroscopic time, such as the time required to measure a current, as van Kampen argues, but rather is some time less than the microscopic relaxation time in which the charges redistribute themselves on the conductor and counterbalance the applied field. Thus, van Kampen's criterion (11.14) is seen as a restriction on the time for which the transient response is linear, rather than a restriction on the strength of fields for which linear response is valid at some time t. It is now seen that van Kampen's criterion and the time dependent response theory have nothing to do with steady state phenomena.

It is possible, however, to deal with steady state responses by assuming that the external field is switched on infinitely slowly. If the system is observed after the interaction reaches full strength, then it is expected that all transient phenomena will have died out and that only the steady state currents will remain. Since the perturbation will have acted on the system with a finite strength for an infinite time, the perturbed and unperturbed systems will have evolved to entirely different states and so no finite number of terms in (11.9) can completely determine $\rho(t)$. The assertion that the actual linear steady state response can be determined from linear response theory depends critically upon the fact that the final state is time-independent. This can most easily be seen by switching on the interaction adiabatically, but it would be expected that the steady state response would be independent of the particular manner in which the interaction was turned on.

This situation may be regarded as analogous to the situation in elementary perturbation theory in which it is often possible to calculate time independent perturbed quantities even though there is no hope of finding the exact time dependent evolution except for very short times. For example, in quantum systems subjected to a periodic disturbance, Fermi's Golden Rule predicts constant transition rates long after the perturbation is switched on, analogous to the steady currents of linear response.

The distinction between time dependent and steady state phenomena may be clarified by considering the general linear response expression for the response of an operator B to a perturbation $-\mu F(t)A$. The statistical average $\langle B(t)\rangle$ of B is given by

$$\begin{aligned}\langle B(t)\rangle &= \mathrm{Tr}[\rho(t)B] \\ &= \mathrm{Tr}[\rho_o B]-\left(\frac{\mu}{ih}\right)\mathrm{Tr}\Big[\int_{-\infty}^{t} dt_1 e^{-i\frac{H_o t}{h}} F(t_1)[A_I(t_1),\rho_o]e^{i\frac{H_o t}{h}}\Big] \, .\end{aligned} \tag{11.15}$$

This may be written, defining $\langle B\rangle_o = \mathrm{Tr}[\rho_o B]$ and using the cyclic invariance of the trace, as

$$\langle B(t)\rangle = \langle B\rangle_o - \left(\frac{\mu}{ih}\right) \int_{-\infty}^{t} dt_1 F(t_1)\mathrm{Tr}\{[A,\rho_o]B_I(t-t_1)\} \, . \tag{11.16}$$

To study transient response, let F(t) be a step function applied at $t=t_o$. Then,

$$\begin{aligned}\langle B(t)\rangle &= \langle B\rangle_o -\left(\frac{\mu}{ih}\right) \int_{t_o}^{t} dt_1 \mathrm{Tr}\{[A,\rho_o]B_I(t-t_1)\} \\ &= \langle B\rangle_o -\left(\frac{\mu}{ih}\right) \int_{o}^{t-t_o} d\tau \mathrm{Tr}\{[A,\rho_o]B_I(\tau)\} \\ &= \langle B\rangle_o + \mu\sigma_1'(t-t_o) \quad .\end{aligned} \tag{11.17}$$

The corresponding non-linear expansion may be written as

$$\langle B(t)\rangle = \langle B\rangle_o + \mu\sigma_1'(t-t_o) + \mu^2\sigma_2'(t-t_o) + \ldots \tag{11.18}$$

where all the coefficients $\sigma_i'(t-t_o)$ are functions of time, but are independent of μ. For this transient response, the non linear terms may be large or may average to a negligible quantity for different times. In this case, the only estimate for the validity of response theory is one in the spirit of van Kampen.

The situation is quite different in the case of adiabatic switching. Adiabatic switching means that the interaction is turned on infinitely slowly so that

$$F(t) = \lim_{\varepsilon\to o^+} e^{\varepsilon t} \quad . \tag{11.19}$$

Thus,

$$\begin{aligned}
\langle B(t)\rangle &= \langle B\rangle_o + \lim_{\varepsilon\to o^+} \left(-\frac{\mu}{i\hbar}\right)\int_{-\infty}^{t} dt_1 e^{\varepsilon t_1} \mathrm{Tr}\{[A,\rho_o]B_I(t-t_1)\} \\
&= \langle B\rangle_o + \lim_{\varepsilon\to o^+}\left(-\frac{\mu}{i\hbar}\right)e^{\varepsilon t}\int_{o}^{\infty} d\tau\, \mathrm{Tr}\{[A,\rho_o]B_I(\tau)e^{-\varepsilon\tau}\} \\
&= \langle B\rangle_o + \mu\sigma_1 \quad .
\end{aligned} \tag{11.20}$$

where

$$\sigma_1 = -\frac{1}{i\hbar}\lim_{\varepsilon\to o^+} e^{\varepsilon t}\int_{o}^{\infty} d\tau\, \mathrm{Tr}\{[A,\rho_o]B_I(\tau)e^{-\varepsilon\tau}\} \tag{11.21}$$

is independent of t for finite times due to the assumption of adiabatic switching. The non-linear terms may be written as

$$\langle B(t)\rangle = \langle B\rangle_o + \mu\sigma_1 + \mu^2\sigma_2 + \ldots \tag{11.22}$$

where all coefficients are independent of t and of μ and depend, for a given perturbation operator A, only upon the properties of the unperturbed system. Thus, the adiabatic switching has resulted in a unique, time independent power series in μ for the steady state response of the system. This power series is exact for any value of μ and the non-linearity as a function of μ is merely a measure of the convergence of the series and not of the validity of a particular term of the expansion.

This, of course, does not give any criterion for the range of applied fields for which the linear response dominates the expansion (11.22) and this would depend upon the dynamics of the system. Since eqn. (11.22) is exact for steady states, it must agree with the true response. Thus, for steady states, the range of validity of linear response theory is the same as the range of validity of the linear phenomenological laws. However, there still remains the question of whether the true phenomenological laws are the ones actually observed in the laboratory. It is possible to imagine some sort of quasi-linearity holding only for reasonably large and measurable values of μ and with true linearity, as described by linear response theory, holding only for values of μ so small that the resulting currents are unmeasurable. This is similar to the possibility discussed in §3.2 with regard to the Onsager reciprocal relations and, as in that discussion, a rigorous justification of the equivalence of linear response expressions and the linear phenomenological laws is missing. However, the linear nature of the laws within the measured range makes such an equivalence reasonable and, for particular models, the range of validity might be calculated.

In summary, it appears that van Kampen's arguments against linear response theory, although valid for transient phenomena, do not affect the validity of the steady state expressions. Instead, investigations as to the validity of steady state expressions should concentrate on the extent to which thermal interactions can be ignored in the physical system and upon whether the observed behavior is truly linear or merely quasi-linear. The non-linear response coefficients are on a much less secure basis since they are directly affected by the lowest order thermal perturbations.

§11.3 Applications

As stated before, linear response theory can be used to derive Onsager relations and fluctuation dissipation theorems. To do this, it is useful to define certain operators and to derive certain relationships.

The perturbing Hamiltonian will be assumed to have M separate contributions and will be written as

$$H_t' = - \sum_{j=1}^{M} F_j(t)A_j \tag{11.23}$$

where the A_j are the Schrödinger operators. The linear response expression can then be written as

$$\rho(t) = \rho_o - \frac{1}{i\hbar} \sum_{j=1}^{M} \int_{-\infty}^{t} F_j(t_1) e^{-i\frac{H_o t}{\hbar}} [A_{jI}(t_1), \rho_o] e^{i\frac{H_o t}{\hbar}} dt_1 \; . \tag{11.24}$$

The linear response of a general operator B_i can be written as

$$\Delta B_i(t) \equiv \langle B_i(t)\rangle - \langle B_i\rangle_o$$

$$= \sum_{j=1}^{M} \left(-\frac{1}{i\hbar}\right) \int_{-\infty}^{t} F_j(t_1) \mathrm{Tr}\{[A_j, \rho_o] B_{iI}(t-t_1)\} dt_1 \quad . \tag{11.25}$$

Define a response function $\phi_{ij}^{BA}(t)$ by

$$\phi_{ij}^{BA}(t) = -\frac{1}{i\hbar} \mathrm{Tr}\{[A_j, \rho_o] B_{iI}(t)\}$$

$$= -\frac{1}{i\hbar} \mathrm{Tr}\{\rho_o [B_{iI}(t), A_j]\} \tag{11.26}$$

and a susceptibility $\chi_{ij}^{BA}(\omega)$ by

$$\chi_{ij}^{BA}(\omega) = \int_{o}^{\infty} \phi_{ij}^{BA}(t_1) e^{-i\omega t_1 - \varepsilon t_1} dt_1 \tag{11.27}$$

where the factor $e^{-\varepsilon t}$ serves to ensure that the integral converges after which ε is allowed to approach o^+.

Then,

$$\Delta B_i(t) = \sum_{j=1}^{M} \int_{-\infty}^{t} \phi_{ij}^{BA}(t-t_1) F_j(t_1) dt_1 \tag{11.28}$$

and, for a periodic perturbation,

$$H'_t = - \sum_{j=1}^{M} F_{jo}\ \mathrm{Re}(e^{i\omega t}) \quad , \tag{11.29}$$

the response is

$$\Delta B_i(t) = \mathrm{Re}\{ \sum_{j=1}^{M} \chi_{ij}^{BA}(\omega) F_{jo}\ e^{i\omega t}\} \quad . \tag{11.30}$$

The ϕ_{ij}^{BA} have certain symmetry properties. First, note that

$$\begin{aligned} \phi_{ij}^{BA*}(t) &= \frac{1}{i\hbar} \mathrm{Tr}\{B_{iI}^{\dagger}(t)[\rho_o^{\dagger},A_j^{\dagger}]\} \\ &= - \frac{1}{i\hbar} \mathrm{Tr}\{[A_j,\rho_o]B_{iI}(t)\} \\ &= \phi_{ij}^{BA}(t) \end{aligned} \tag{11.31}$$

where * denotes complex conjugate, † denotes Hermitian conjugate, and where A_j, B_i have been assumed to be Hermitian.

Also, note that, if time and all magnetic fields $\vec{H}$, are reversed, the wave functions and operators in the position representation go to their complex conjugates. Thus,

$$\begin{aligned} \phi_{ij}^{BA}(-t,-\vec{H}) &= \frac{1}{i\hbar} \mathrm{Tr}^*\{\rho_o^*[B_{iI}(t)^*,A_j^*]\} \\ &= - \varepsilon_i\varepsilon_j\ \phi_{ij}^{BA*}(t,\vec{H}) \\ &= - \varepsilon_i\varepsilon_j\ \phi_{ij}^{BA}(t,\vec{H}) \end{aligned} \tag{11.32}$$

where $\varepsilon_i, \varepsilon_j = \pm 1$ depending on whether B_i, A_j are even or odd under time reversal.

Finally, note that

$$\phi_{ij}^{BA}(-t) = -\frac{1}{i\hbar}\,\mathrm{Tr}\{\rho_o[B_{iI}(-t), A_j]\}$$

$$= \frac{1}{i\hbar}\,\mathrm{Tr}\{\rho_o[A_{jI}(t), B_i]\}$$

$$= -\,\phi_{ji}^{AB}(t) \quad . \tag{11.33}$$

These symmetry relations are most useful in the case where $B_i = A_i$. Then, letting $\phi_{ij} = \phi_{ij}^{AA}$,

$$\phi_{ij}^{*}(t,\vec{H}) = \phi_{ij}(t,\vec{H}) \tag{11.34}$$

and

$$\phi_{ij}(t,\vec{H}) = \varepsilon_i\varepsilon_j\phi_{ji}(t,-\vec{H}) \quad . \tag{11.35}$$

Eqn. (11.35) can be used to derive Onsager relations in the case where the currents are the time derivatives of the "forces" A_i. That is, if

$$J_i = \frac{dA_i}{dt} \tag{11.36}$$

then

$$J_i = \sum_{j=1}^{M} \left[\int_{-\infty}^{t} F_j(t_1)\,\frac{\partial\phi_{ij}(t-t_1,\vec{H})}{\partial t}\,dt_1 + F_j(t)\phi_{ij}(o)\right] \quad . \tag{11.37}$$

The second term in (11.37) is zero if it is assumed that $[A_i, A_j] = 0$, as is often the case. Then, for F_j given by eqn. (11.29),

$$J_i(\omega) = \sum_{j=1}^{M} \mathrm{Re}\{\int_{-\infty}^{t} F_{jo} e^{i\omega t_1} \frac{\partial\phi_{ij}(t-t_1,\vec{H})}{\partial t} dt_1\}$$

$$= \sum_{j=1}^{M} \mathrm{Re}\{F_{jo} e^{i\omega t} \int_{0}^{\infty} e^{-i\omega\tau} \frac{\partial\phi_{ij}(\tau,\vec{H})}{\partial\tau} d\tau\}$$

$$= \sum_{j=1}^{M} \varepsilon_i\varepsilon_j \mathrm{Re}\{F_{jo} e^{i\omega t} \int_{0}^{\infty} e^{-i\omega\tau} \frac{\partial\phi_{ji}(\tau,-\vec{H})}{\partial\tau} d\tau\} \quad . \qquad (11.38)$$

Defining $L_{ij}(\omega,\vec{H})$ by

$$L_{ij}(\omega,\vec{H}) = \int_{0}^{\infty} e^{-i\omega\tau} \frac{\partial\phi_{ij}(\tau,\vec{H})}{\partial\tau} d\tau \qquad (11.39)$$

then

$$J_i = \sum_{j=1}^{M} \mathrm{Re}\{L_{ij}(\omega,\vec{H}) F_{jo} e^{i\omega t}\} \qquad (11.40)$$

and the L_{ij} satisfy the symmetry relations

$$\varepsilon_i\varepsilon_j L_{ij}(\omega,\vec{H}) = L_{ji}(\omega,-\vec{H}) \quad . \qquad (11.41)$$

These are the Onsager relations. They are based on several assumptions which are similar to those required for Casimir's proof, as presented in §3.2. First, it is assumed that the linear response expressions are the same as those observed phenomenologically (i.e. that no quasi-linearity holds). Secondly, it was assumed that the currents considered are the time derivatives of the "forces". Also, in accordance with the earlier discussion of linear response theory itself, the system must have a relaxation time much less than the time in which the perturbation is turned on. Finally, in order that the linear response theory should be a good approximation to the full response, the system must be "near" equilibrium. It is interesting that the assumptions needed to derive Onsager relations from linear response theory are almost identical to those needed for Casimir's proof. Thus, linear response merely provides an alternative derivation of Onsager relations which does not directly involve thermodynamic concepts such as entropy and in which the role of fluctuations is hidden in the density operator ρ_0.

As an example, consider the case of electrical conduction. Then, H'_t may be written as

$$H'_t = \sum_{\mu=1}^{N} e\,\vec{r}_\mu \cdot \vec{E}(t)$$

$$= -e \sum_{\mu=1}^{N} \sum_{i=1}^{3} r_{\mu i}\, E_i \qquad (11.42)$$

where $\vec{r}_\mu$ is the position of the μ^{th} particle and where e is the charge on each of the N particles. Then, identifying $A_i = e \sum_{\mu=1}^{N} r_{\mu i}$, the current J_i in the i^{th} direction is given by

$$J_i = e \sum_{\mu=1}^{N} \dot{r}_{\mu i} = \frac{d(e \sum_{\mu=1}^{N} r_{\mu i})}{dt} = \frac{dA_i}{dt} \quad . \qquad (11.43)$$

Therefore, from the general results above,

$$J_i = \sum_{j=1}^{3} \mathrm{Re}[\sigma_{ij}(\omega,\vec{H})E_j(t)] \qquad (11.44)$$

for periodic electric fields. In (11.43), L_{ij} has been written as σ_{ij} in accordance with common usage. The conductivity tensor σ_{ij} obeys the Onsager relation

$$\sigma_{ij}(\omega,\vec{H}) = \sigma_{ji}(\omega,-\vec{H}) \qquad (11.45)$$

and is given by the expression

$$\sigma_{ij}(\omega,\vec{H}) = \int_0^\infty e^{-i\omega\tau}\left(-\frac{1}{i\hbar}\right)\mathrm{Tr}\{[A_j,\rho_o]\, J_{iI}(\tau)\}d\tau \quad . \qquad (11.46)$$

For the canonical ensemble, $\rho_o \propto e^{-\beta H_o}$, this can be written in a compact form by using the identity

$$[A_j,e^{-\beta H_o}] = e^{-\beta H_o}\int_0^\beta e^{\lambda H_o}[H_o,A_j]e^{-\lambda H_o}d\lambda$$

$$= -i\hbar\, e^{-\beta H_o}\int_0^\beta \dot{A}_{ji}(-i\hbar\lambda)d\lambda \qquad (11.47)$$

where the dot denotes time derivative (cf. Kubo [44] for a derivation of this identity). For the canonical ensemble,

$$\sigma_{ij}(\omega,\vec{H}) = \int_0^\infty e^{-i\omega\tau}\int_0^\beta \mathrm{Tr}\{\rho_o J_{jI}(-i\hbar\lambda)J_{iI}(t)\}d\lambda\, d\tau \qquad (11.48)$$

which is usually written simply as

$$\sigma_{ij}(\omega,\vec{H}) = \int_0^\infty e^{-i\omega\tau}\int_0^\beta \langle J_j(-i\hbar\lambda)J_i(t)\rangle_o d\lambda\, d\tau \quad . \qquad (11.49)$$

The form (11.49) is usually referred to as a Kubo formula.

As a second application of linear response theory, a fluctuation dissipation theorem is derived. In its conventional form, a fluctuation dissipation theorem establishes a relationship between the linear response $\Delta q(t)$ of a system to an external perturbation and the equilibrium time correlations $<q(o)q(t)>_0$ of the same quantity, q. It is, however, more convenient to express this as a relationship between the frequency dependent response function $\chi_q(\omega)$ and the spectral density of fluctuations $S_q(\omega)$. These quantities are defined by

$$\Delta q(t) = \mathrm{Re}\{\chi_q(\omega)F_0\, e^{i\omega t}\} \quad , \qquad (11.50)$$

for perturbations of the form (11.29), and by

$$S_q(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} <q(o)q(t)>_0\, e^{i\omega t}dt \quad . \qquad (11.51)$$

The original fluctuation dissipation theorem was obtained by Nyquist [47] in the case of an electrical network. Later, his formula was generalized to other systems and a formula was obtained which reduced to his formula in the classical limit.

In order to derive this fluctuation dissipation formula from linear response theory, define the analogues of χ_q and $S_q(\omega)$ by

$$\chi_Q = \int_0^{\infty} \left(-\frac{1}{i\hbar}\right)\mathrm{Tr}\{[Q,\rho_0]Q_I(t)\}e^{-i\omega t}dt \qquad (11.52)$$

where Q is the quantum operator of the perturbation whose response is described by χ_Q. Then,

$$\Delta Q(t) = \mathrm{Re}\{\chi_Q(\omega) F_o\, e^{i\omega t}\} \quad . \tag{11.53}$$

The quantum equivalent of the autocorrelation function $\langle q(o)q(t)\rangle_o$ is the properly symmetrized product of Q and $Q_I(t)$. That is, consider

$$\tfrac{1}{2}\langle Q_I(o)Q_I(t) + Q_I(t)Q_I(o)\rangle_o = \tfrac{1}{2}\langle QQ_I(t) + Q_I(t)Q\rangle_o \tag{11.54}$$

and define the fluctuation spectrum

$$S_Q(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} \tfrac{1}{2}\langle QQ_I(t) + Q_I(t)Q\rangle_o e^{i\omega t}dt \quad . \tag{11.55}$$

Using a canonical ρ_o and the eigenstates of H_o as a basis,

$$\tfrac{1}{2}\langle QQ_I(t) + Q_I(t)Q\rangle_o = \tfrac{1}{2}\mathrm{Tr}\{\rho_o(QQ_I(t) + Q_I(t)Q)\}$$

$$= \tfrac{1}{2}[\mathrm{Tr}\rho_o]^{-1} \sum_{k,\ell} e^{-\beta E_k} Q_{k\ell}Q_{\ell k}\{e^{it(E_\ell - E_k)} + e^{-it(E_\ell - E_k)}\}$$

$$\cdot \tag{11.56}$$

where $\hbar$ has been set equal to one for convenience. Hence,

$$S_Q(\omega) = [\mathrm{Tr}\rho_o]^{-1} \sum_{k,\ell} e^{-\beta E_k}|Q_{k\ell}|^2\{\delta(E_k - E_\ell + \omega) + \delta(E_k - E_\ell - \omega)\}$$

$$= [\mathrm{Tr}\rho_o]^{-1}(1 + e^{-\beta\omega}) \sum_{k,\ell} e^{-\beta E_k}|Q_{k\ell}|^2 \delta(E_k - E_\ell + \omega) \quad ; \tag{11.57}$$

Calculating χ_Q from (11.52), in the same basis, yields

$$\mathrm{Im}(\chi_Q(\omega)) = -\pi[\mathrm{Tr}\rho_o]^{-1}(1-e^{-\beta\omega})\sum_{k,\ell} e^{-\beta E_k}|Q_{k\ell}|^2\delta(E_k-E_\ell+\omega) \quad . \tag{11.58}$$

Comparing equations (11.58) and (11.57) gives the fluctuation dissipation theorem as derived by Callen and Welton [48]:

$$S_Q(\omega) = -\frac{1}{\pi}\frac{(1+e^{-\beta\omega})}{(1-e^{-\beta\omega})}\mathrm{Im}(\chi_Q(\omega)) \tag{11.59}$$

which in the classical (high temperature) limit, $\beta\omega \to 0$, gives Nyquist's formula

$$S_Q(\omega) = -\frac{2}{\pi}\frac{1}{\beta\omega}\mathrm{Im}(\chi_Q(\omega)) \quad . \tag{11.60}$$

This concludes the discussion of linear response theory. It should be kept in mind that, although various authors have derived linear response expressions for non-mechanical perturbations (e.g. diffusion transport coefficients), linear response does not provide a general method for finding effects due to non-mechanical interactions. Also, note that no general procedure is available for calculating the linear response expressions and that the actual calculations can only be carried out for some drastically simplified form of H_o.

CHAPTER XII

PRIGOGINE THEORY OF IRREVERSIBLE PROCESSES

§12.1 Introduction and Potential Scattering

In a series of papers, Prigogine and his collaborators developed a rather different approach to non-equilibrium statistical mechanics (Prigogine and Balescu [49] and [50], also Prigogine [51]). Starting from the Liouville equation, they obtained equations for the Fourier components of the distribution function. These Fourier components have direct interpretations as correlations in the system. Also, the Fourier expansion is formally an expansion in the eigenfunctions of the free particle Liouville operator. There is a direct analogy to the Schrödinger equation in quantum mechanics and many of the techniques used in quantum mechanics may be applied to the Liouville equation. By putting the Hamiltonian in the form $H = H_0 + \lambda V$ where H_0 is the free particle Hamiltonian and λV describes the interactions, it is possible to associate "matrix elements" with the Liouville operator which give rise to transitions between different eigenstates of H_0. These are states of different correlation and thus the Prigogine theory is described as a theory of the dynamics of correlations. Also, this formalism is conveniently written in terms of of diagrams which represent the different possible transitions and which are referred to as representing the creation, destruction, or propagation of correlations. By concentrating on "dominant diagrams" only, it is sometimes possible to derive closed equations for certain of the Fourier terms and to relate these to the usual reduced distribution functions. In choosing dominant diagrams, an important role is played by the limit $N\to\infty$, Ω(volume)$\to\infty$ with N/Ω finite and by the limit $\lambda\to 0$, $t\to\infty$. Finally, much of the theory is most clearly and concisely formulated by using a resolvent technique to obtain iterative equations for the correlations.

Most of the specific equations derived from the Prigogine theory are obtainable from other formalisms. This is a reflection of the fact that the complete equations of Prigogine are entirely equivalent to other schemes such as the BBGKY theory and Prigogine's choice of dominant terms is related to the approximations made in other theories. However, as emphasized by Bogolyubov, correlations play an extremely important role in irreversibility and in the study of kinetic equations. Thus, since the theory is explicitly set up to deal with correlations, it often gives a simpler formulation of assumptions and approximations than is possible in other theories. The generalization of the results of the Prigogine theory to quantum mechanics has been carried out and presents no essential difficulty (cf. Resibois [52] or Prigogine [51]). However, only the classical case will be treated in this thesis.

For the classical case, the system's evolution is governed by the Liouville equation for the full distribution function $\rho(\vec{X},t)$:

$$\frac{\partial\rho(\vec{X},t)}{\partial t} = \{H,\rho\} \tag{12.1}$$

where $\vec{X}$ represents the 6N coordinates of the system. Using the Liouville operator L, defined in §9.3 as

$$L = i\{H,\} \quad , \tag{12.2}$$

eqn. (12.1) becomes

$$\frac{\partial\rho}{\partial t} = -i\, L\, \rho \quad . \tag{12.3}$$

Since L is linear and Hermitian, eqn. (12.3) is analogous to the Schrödinger equation

$$\frac{\partial\psi}{\partial t} = -\frac{i}{\hbar} H\, \psi \tag{12.4}$$

so the methods of quantum mechanics are appropriate.

First, consider a single free particle ($H = H_o = \frac{p^2}{2m}$). Then,

$$L = L_o = -i\, \frac{\vec{p}}{m} \cdot \frac{\partial}{\partial\vec{r}} \tag{12.5}$$

so eqn. (12.1) is just

$$\frac{\partial\rho}{\partial t} = -i\, L_o\, \rho = -\frac{\vec{p}}{m} \cdot \frac{\partial\rho}{\partial\vec{r}} \quad . \tag{12.6}$$

The eigenfunctions of L_o are the functions $\phi_{\vec{k}}$ defined by

$$L_o\, \phi_{\vec{k}} = \lambda_{\vec{k}}\, \phi_{\vec{k}} \quad . \tag{12.7}$$

These are easily seen to be

$$\phi_{\vec{k}} = \Omega^{-1/2}\, e^{i\vec{k}\cdot\vec{r}}\, \rho_{\vec{k}}(\vec{p}) \tag{12.8}$$

corresponding to the real eigenvalues

$$\lambda_{\vec{k}} = \frac{\vec{p}}{m} \cdot \vec{k} \quad . \tag{12.9}$$

The coefficients $\rho_{\vec{k}}(\vec{p})$ are normalized so that

$$\int |\rho_{\vec{k}}(\vec{p})|^2\, d^3\vec{p} = 1 \quad . \tag{12.10}$$

With these definitions,

$$\rho(\vec{r},\vec{p},t) = \Omega^{-1/2} \sum_{\vec{k}} e^{-i\lambda_{\vec{k}} t}\, \phi_{\vec{k}}(\vec{r},\vec{p})$$

$$= \Omega^{-1} \sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{r}-\frac{\vec{p}t}{m})}\, \rho_{\vec{k}}(\vec{p}) \quad . \tag{12.11}$$

The coefficients $\rho_{\vec{k}}(\vec{p})$ are to be determined by the initial condition on $\rho(\vec{r},\vec{p},t)$ so

$$\rho(\vec{r},\vec{p},o) = \Omega^{-1} \sum_{\vec{k}} e^{i\vec{k}\cdot\vec{r}}\, \rho_{\vec{k}}(\vec{p}) \quad . \tag{12.12}$$

If there is an interaction term λV in the Hamiltonian, then the distribution function may still be expanded as

$$\rho(\vec{r},\vec{p},t) = \Omega^{-1} \sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{r}-\frac{\vec{p}t}{m})} \rho_{\vec{k}}(\vec{p},t) \qquad (12.13)$$

but the coefficients $\rho_{\vec{k}}(\vec{p},t)$ are no longer independent of time. In this case, the Liouville operator may be written as

$$L = L_0 + \lambda \delta L \qquad (12.14)$$

so the Liouville equation becomes

$$\frac{\partial\rho}{\partial t} = -i(L_0 + \lambda\delta L)\rho \quad . \qquad (12.15)$$

Inserting the expansion (12.13) for ρ gives

$$i\Omega^{-1} \sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{r}-\frac{\vec{p}t}{m})}(-i\vec{k}\cdot\frac{\vec{p}}{m})\rho_{\vec{k}}(\vec{p},t) + i\Omega^{-1} \sum_{\vec{k}} e^{i\vec{k}\cdot(\vec{r}-\frac{\vec{p}}{m}t)}\frac{\partial\rho_{\vec{k}}}{\partial t}$$

$$= \Omega^{-1} \sum_{\vec{k}} (\frac{\vec{p}\cdot\vec{k}}{m}) e^{i\vec{k}\cdot(\vec{r}-\frac{\vec{p}t}{m})} \rho_{\vec{k}}(\vec{p},t) + \Omega^{-1} \sum_{\vec{k}} \delta L e^{i\vec{k}\cdot(\vec{r}-\frac{\vec{p}t}{m})} \rho_{\vec{k}}(\vec{p},t). \qquad (12.16)$$

Multiplying by $e^{i\vec{k}'\cdot\vec{r}}$ and integrating over $\vec{r}$ yields

$$i\frac{\partial\rho_k}{\partial t} = \lambda \sum_{\vec{k}} e^{i\vec{k}\cdot\frac{\vec{p}}{m}t} \langle k|\delta L|k'\rangle e^{-i\vec{k}'\cdot\frac{\vec{p}}{m}t} \rho_{\vec{k}'}(\vec{p},t) \qquad (12.17)$$

where the "matrix elements" are defined by

$$\langle k|\delta L|k'\rangle = \Omega^{-1} \int e^{-i\vec{k}\cdot\vec{r}} \delta L e^{i\vec{k}'\cdot\vec{r}} d^3\vec{r} \quad . \qquad (12.18)$$

It should be kept in mind that the matrix elements are operators.

The initial values of the $\rho_{\vec{k}}$ are given by

$$\rho_{\vec{k}}(\vec{p},o) = \int e^{-i\vec{k}\cdot\vec{r}}\rho(\vec{r},\vec{p},o)d^3\vec{r} \quad . \tag{12.19}$$

Before generalizing eqn. (12.17) to the case of more than one particle, consider an example where λV represents the effect of an external central potential $V(|\vec{r}|)$ on a free particle. Then,

$$\delta L = i\{V,\} = i\,\frac{\partial V}{\partial\vec{r}}\cdot\frac{\partial}{\partial\vec{p}} \tag{12.20}$$

so

$$<\vec{k}|\delta L|\vec{k}'> = i\Omega^{-1}\int e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}\,\frac{\partial V}{\partial\vec{r}}d^3\vec{r}\cdot\frac{\partial}{\partial\vec{p}} \quad . \tag{12.21}$$

Fourier expanding V as

$$V(|\vec{r}|) = \Omega^{-1}\sum_{\vec{\ell}} V_{\vec{\ell}}\, e^{i\vec{\ell}\cdot\vec{r}} \tag{12.22}$$

where

$$V_{\vec{\ell}} = \int e^{-i\vec{\ell}\cdot\vec{r}}V(|\vec{r}|)d^3\vec{r} \tag{12.23}$$

gives

$$<\vec{k}|\delta L|\vec{k}'> = -\Omega^{-1}\sum_{\vec{\ell}} V_{\vec{\ell}}\,\delta_{\vec{k}',\vec{k}-\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}} \quad . \tag{12.24}$$

Eqn. (12.17) now becomes

$$i\frac{\partial\rho_{\vec{k}}}{\partial t} = \lambda\Omega^{-1}\sum_{\vec{\ell}} e^{i\vec{k}\cdot\frac{\vec{p}}{m}t}\,V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\left[e^{-i(\vec{k}+\vec{\ell})\cdot\frac{\vec{p}}{m}t}\rho_{\vec{k}+\vec{\ell}}(\vec{p},t)\right] \tag{12.25}$$

where the fact has been used that $V_{\vec{\ell}} = V_{-\vec{\ell}}$ for a central potential.

Eqn. (12.25) can be rewritten as an integral equation

$$\rho_{\vec{k}}(\vec{p},t) = \rho_{\vec{k}}(\vec{p},o)+\left(\frac{\lambda}{i\Omega}\right)\int_o^t \sum_{\vec{\ell}} e^{i\frac{(\vec{k}\cdot\vec{p})}{m}t_1} V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\left[e^{-i(\vec{k}+\vec{\ell})\cdot\frac{p}{m}t_1}\rho_{\vec{k}+\vec{\ell}}(\vec{p},t_1)\right]dt_1 \tag{12.26}$$

which can be iterated to yield

$$\begin{aligned}\rho_{\vec{k}}(\vec{p},t) = {} & \rho_{\vec{k}}(\vec{p},o)+\left(\frac{\lambda}{i\Omega}\right)\int_o^t \sum_{\vec{\ell}} e^{i\frac{(\vec{k}\cdot\vec{p})}{m}t_1} V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\left[e^{-i(\vec{k}+\vec{\ell})\cdot\frac{\vec{p}}{m}t_1}\rho_{\vec{k}+\vec{\ell}}(\vec{p},o)\right]dt_1 \\ & +\left(\frac{\lambda}{i\Omega}\right)^2\int_o^t\int_o^{t_1}\sum_{\vec{\ell},\vec{\ell}'} e^{i\frac{(\vec{k}\cdot\vec{p})}{m}t_1} V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\Big[e^{-i(\vec{k}+\vec{\ell})\cdot\frac{\vec{p}}{m}t_1} e^{i(\vec{k}+\vec{\ell})\cdot\frac{\vec{p}}{m}t_2} \\ & V_{\vec{\ell}'}\,\vec{\ell}'\cdot\frac{\partial}{\partial\vec{p}}\left[e^{-i(\vec{k}+\vec{\ell}+\vec{\ell}')\cdot\frac{\vec{p}}{m}t_2}\rho_{\vec{k}+\vec{\ell}+\vec{\ell}'}(\vec{p},o)\right]\Big]dt_1\,dt_2 + \ldots \quad .\end{aligned} \tag{12.27}$$

This equation describes the evolution of all the Fourier components. Often, however, only the total momentum distribution is of interest and not the local inhomogeneities. For example, one might be interested in the total number of particles scattered into a particular final velocity state. In that case, the quantity of interest is the momentum distribution

$$f(\vec{p},t) = \int_{\Omega} \rho(\vec{r},\vec{p},t)\,d^3\vec{r} \quad . \tag{12.28}$$

Thus, it is sufficient to study the evolution of ρ_o alone. The relevant equation is

$$\rho_o(\vec{p},t) = \rho_o(\vec{p},o)+(\frac{\lambda}{i\Omega})\int_o^t \sum_{\vec{\ell}} V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}[e^{-i\frac{\vec{\ell}\cdot\vec{p}}{m}t_1}\,\rho_{\vec{\ell}}(\vec{p},o)]dt_1$$

$$+(\frac{\lambda}{i\Omega})^2\int_o^t\int_o^{t_1} \sum_{\vec{\ell},\vec{\ell}'} V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}[e^{i\frac{\vec{\ell}\cdot\vec{p}}{m}(t_2-t_1)}\,V_{\vec{\ell}'}\,\vec{\ell}'\cdot\frac{\partial}{\partial\vec{p}}[e^{-i(\vec{\ell}+\vec{\ell}')\cdot\frac{\vec{p}}{m}t_2}$$

$$\rho_{\vec{\ell}+\vec{\ell}'}(\vec{p},o)]]dt_1dt_2$$

$$+\ldots \quad . \tag{12.29}$$

If it is assumed that there are no initial correlations, then

$$\rho(\vec{r},\vec{p},o) = \rho_o(\vec{p},o)$$

and

$$\rho_o(\vec{p},t) = \rho_o(\vec{p},o)+(\frac{\lambda}{\Omega})^2\int_o^t\int_o^{t_1} \sum_{\vec{\ell}} V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}[e^{i\frac{\vec{\ell}\cdot\vec{p}}{m}(t_2-t_1)}\,V_{\vec{\ell}}\,\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\rho_o(\vec{p},o)]$$

$$dt_1dt_2 + O(\lambda^3)$$

$$= \rho_o(\vec{p},o)+(\frac{\lambda}{\Omega})^2\sum_{\vec{\ell}}|V_{\vec{\ell}}|^2\vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\{[\int_o^t\int_o^{t_1} e^{i\frac{\vec{\ell}\cdot\vec{p}}{m}(t_2-t_1)}dt_1dt_2]$$

$$\vec{\ell}\cdot\frac{\partial\rho_o(\vec{p},o)}{\partial\vec{p}}\} + O(\lambda^3). \tag{12.30}$$

To further simplify this, note that an integral of the form

$$I = \int_o^t \int_o^{t_1} e^{i\alpha(t_1-t_2)} dt_1 dt_2 \qquad (12.31)$$

may be integrated, using the variables $\tau = t_1 - t_2$ and $T = t_1 + t_2$, to give

$$I = \int_o^t (t-\tau) e^{i\alpha\tau} d\tau$$

$$= -\frac{t}{i\alpha} + \frac{(e^{i\alpha t}-1)}{(i\alpha)^2} \quad . \qquad (12.32)$$

This expression, with the appropriate value of α, is to be inserted into eqn. (12.30). Noting that $V_{\vec{\ell}} = V_{-\vec{\ell}}$, the first term in eqn. (12.32) gives no contribution in the sum over $\vec{\ell}$ since the function summed over is odd in $\vec{\ell}$. Thus, eqn. (12.30) becomes

$$\rho_o(\vec{p},t) = \rho_o(\vec{p},o) - m^2\left(\frac{\lambda}{\Omega}\right)^2 \sum_{\vec{\ell}} |V_{\vec{\ell}}|^2 \vec{\ell}\cdot\frac{\partial}{\partial\vec{p}}\left\{\frac{[e^{i\frac{\vec{\ell}\cdot\vec{p}}{m}t}-1]}{(\vec{\ell}\cdot\vec{p})^2}\vec{\ell}\cdot\frac{\partial\rho_o(\vec{p},o)}{\partial\vec{p}}\right\}$$

ignoring terms of higher order in λ. (12.33)

From the iterative nature of the derivation of eqn. (12.33) as well as from the discussion of similar expansions in §9.3, it is apparent that this expression (ignoring terms of order higher than λ^2) is valid only for times sufficiently small that the effect of the potential λV upon the distribution function is negligible. Prigogine, however, consistently takes the limit $t\to\infty$ while ignoring higher order terms (cf. Wu [53]). He justifies this procedure by considering the "weak coupling" limit $t \to \infty$, $\lambda \to 0$ with $\lambda^2 t$ constant. This limit, whatever its physical significance, then allows him to use the asymptotic ($t \to \infty$) values of the expressions in eqn. (12.33). He also takes the limit $\Omega \to \infty$ in order to replace the sum over $\vec{\ell}$ by an integral, using the well known correspondence

$$\sum_{\vec{\ell}} \to \frac{\Omega}{(2\pi)^3} \int d\vec{\ell} \quad . \qquad (12.34)$$

In Prigogine's treatment of this problem, the various limits and integrations are exchanged without any justification being given. This, combined with the fact that his results are essentially low order perturbation theory and are critically dependent upon the weak coupling assumption, seems to detract from the rigour and generality claimed for his results. Also, Prigogine's interpretation of his results on the scattering problem (cf. Prigogine [54]) can be questioned (cf. Kreuzer and Kurihara [55]) since it may be argued that a single potential of finite strength in an infinite volume can have only a local effect. Thus, in considering the effect of the potential upon $\rho_0(\vec{p},t)$, which refers to the system as a whole, the effect will be negligible as $\Omega \to \infty$. This is seen in the fact that the correction to $\rho_0(\vec{p},o)$ in eqn. (12.33) is proportional to Ω^{-2}. Without further analysis of the limiting procedures used, it is not obvious that the expression (12.33) gives the dominant contribution for $t \to \infty$, even in the weak coupling limit.

§12.2 General Theory and Resolvent Formalism

The main interest in the Prigogine theory, however, lies not in the interaction of a single particle (or an ideal gas) with an external potential, but rather in analysing the behavior of an interacting system of particles. The above results for a single particle are easily generalized to the case of N particles. The appropriate expansion of ρ is then

$$\rho(\vec{r}_1,\vec{p}_1,\ldots\vec{r}_N,\vec{p}_N,t) = \Omega^{-N} \sum_{\{\vec{k}\}} \rho_{\{\vec{k}\}}(\{\vec{p}\},t)e^{i\sum_{j=1}^{N}\vec{k}_j\cdot(\vec{r}_j-\frac{\vec{p}_j}{m}t)} \tag{12.35}$$

where $\{\vec{k}\} = (\vec{k}_1,\ldots\vec{k}_N)$; $\{\vec{p}\} = (\vec{p}_1,\ldots,\vec{p}_N)$

and where $\sum_{\{\vec{k}\}}$ is over all possible sets $\{\vec{k}\}$. For the general case, assume the Hamiltonian is of the form

$$\begin{aligned} H &= H_o + \lambda V \\ &= \sum_{i=1}^{N} \frac{p_i^2}{2m} + \lambda \sum_{1\leq j<i\leq N} V_{ji}(|\vec{r}_j-\vec{r}_i|) \quad . \end{aligned} \tag{12.36}$$

Then, L is of the form

$$\begin{aligned} L &= L_o + \lambda\delta L \\ &= -i\sum_{i=1}^{N}\frac{\vec{p}_i}{m}\cdot\frac{\partial}{\partial\vec{r}_i} - i\lambda\sum_{1\leq j<i\leq N}\frac{\partial V_{ji}}{\partial\vec{r}_j}\cdot\vec{D}_{ij} \end{aligned} \tag{12.37}$$

where

$$\vec{D}_{ij} = \left(\frac{\partial}{\partial \vec{p}_i} - \frac{\partial}{\partial \vec{p}_j}\right) \quad . \tag{12.38}$$

Manipulations similar to those leading to eqn. (12.17) give the equation of motion for $\rho_{\{\vec{k}\}}$ as

$$i\frac{\partial \rho_{\{\vec{k}\}}(\{\vec{p}\},t)}{\partial t} = \lambda \sum_{\{\vec{k}\}} e^{i\sum_{j=1}^{N} \vec{k}_j \cdot \frac{\vec{p}_j}{m}t} \langle\{\vec{k}\}|\delta L|\{\vec{k}'\}\rangle e^{-i\sum_{j=1}^{N} \vec{k}'_j \cdot \frac{\vec{p}_j}{m}t} \rho_{\{\vec{k}'\}}(\{\vec{p}\},t) \tag{12.39}$$

where the matrix elements are given by

$$\langle\{\vec{k}\}|\delta L|\{\vec{k}'\}\rangle = \Omega^{-1} \sum_{\vec{\ell}} \sum_{1\leq j<i\leq N} V_{\vec{\ell}}\, \vec{\ell}\cdot\vec{D}_{ij} \prod_{n\neq i,j} \delta_{\vec{k}_n,\vec{k}'_n}\, \delta_{\vec{k}'_i,\vec{k}_i-\vec{\ell}}\, \delta_{\vec{k}'_j,\vec{k}_j+\vec{\ell}} \tag{12.40}$$

where the potential $V(|\vec{r}_i-\vec{r}_j|)$ has been Fourier expanded as

$$V(|\vec{r}_i-\vec{r}_j|) = \Omega^{-1} \sum_{\vec{\ell}} V_{\vec{\ell}}\, e^{i\vec{\ell}\cdot(\vec{r}_j-\vec{r}_i)} \quad . \tag{12.41}$$

In a manner entirely analogous to the derivation of eqn. (12.33), a perturbation expression for the components $\rho_{\{\vec{k}\}}(\{\vec{p}\},t)$ may be derived from eqn. (12.40). However, a more concise development of the formalism is possible by introducing resolvent operators. First, however, it is useful to introduce the diagrams corresponding to the different possible transitions between states of different correlations.

Note that the matrix element $\langle\{\vec{k}\}|\delta\mathcal{L}|\{\vec{k}'\}\rangle$ is non-zero only if

$$\sum_{i=1}^{N} \vec{k}_i = \sum_{i=1}^{N} \vec{k}'_i \quad . \qquad (12.42)$$

That is, wave vectors are conserved in the various possible transitions. Thus, a typical matrix element would be

$$\langle\{\vec{k}\},\vec{k}_i,\vec{k}_j|\delta\mathcal{L}|\{\vec{k}\},\vec{k}_i-\vec{\ell},\vec{k}_j+\vec{\ell}\,\rangle \quad . \qquad (12.43)$$

This element may be pictured as in Fig. 12.1a. In standard scattering notation, this would correspond to Fig. 12.1b.

Since the cases of interest are often those in which there are few initial correlations (all or most of the $\vec{k}_i$ are zero), only non-zero wave vectors will be shown in the corresponding diagrams. The state with all $\vec{k}_i$ zero (denoted by $|o\rangle$) is said to form a "correlation vacuum".

There are six basic types of diagrams, shown in Fig. 12.2. Figures (a) and (b) represent the propagation of correlations from the right to the left, (c) and (d) represent the creation of correlations, and (e) and (f) represent the destruction of correlations. All possible matrix elements can be represented as combinations of such "elementary" interaction diagrams.

In order to derive closed equations for the various $\rho_{\{\vec{k}\}}$, Prigogine studied the asymptotic dependence

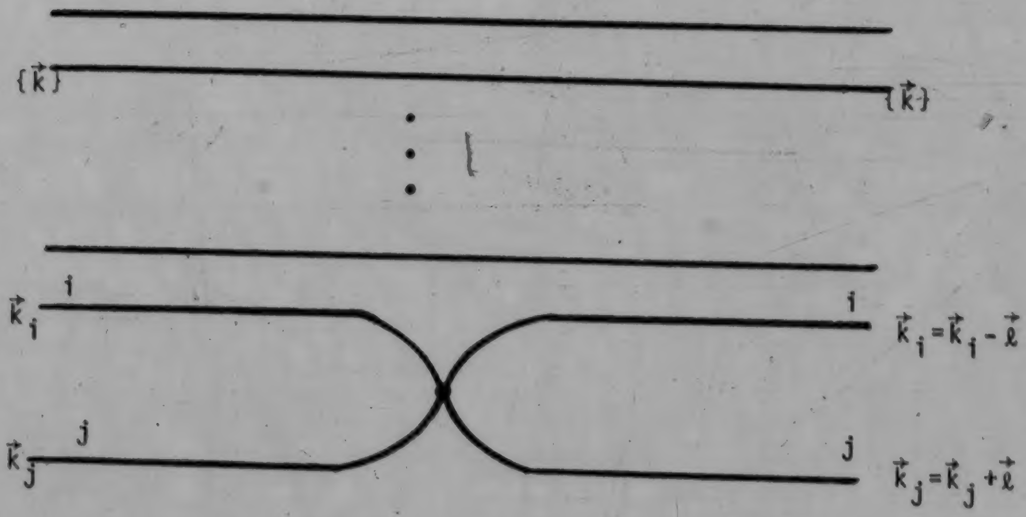

$\{\vec{k}\}$
$\{\vec{k}\}$
$\vec{k}_i$
i
i
$\vec{k}_i=\vec{k}_i-\vec{\ell}$
$\vec{k}_j$
j
j
$\vec{k}_j=\vec{k}_j+\vec{\ell}$

Fig. 12.1a: A typical correlation diagram.

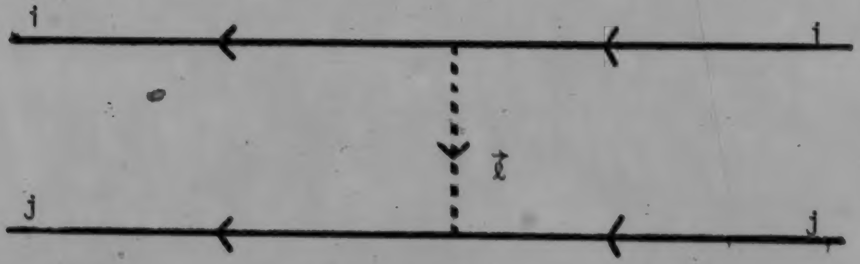
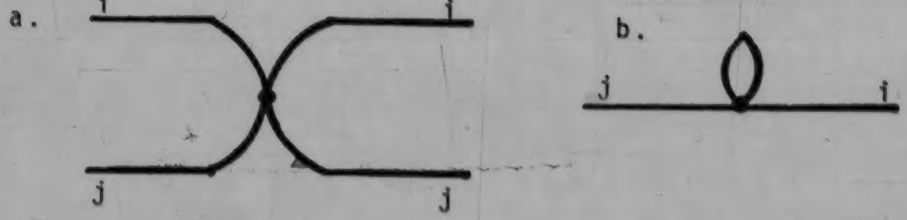
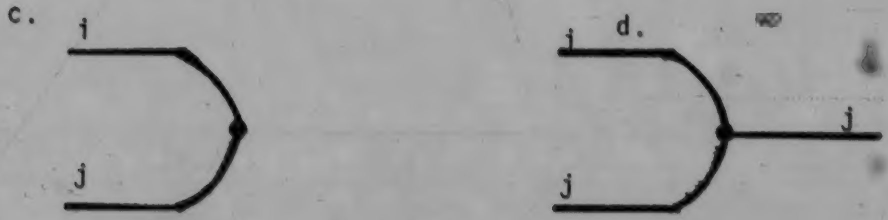

i
i
$\vec{\ell}$
j
j

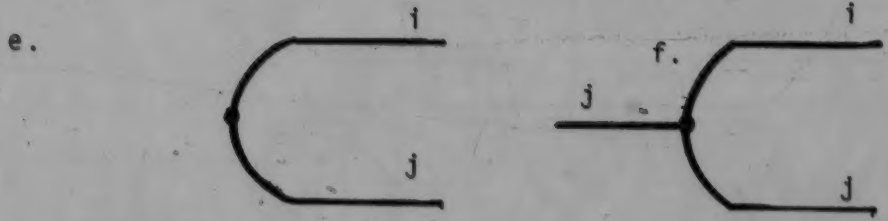
Fig. 12.1b: Above Fig. 12.1a in standard scattering notation.



Propagation



Creation



Destruction

Fig. 12.2: Six "elementary" diagrams

of the various diagrams upon N and Ω (in the limit $N,\Omega\to\infty, N/\Omega$ finite). This required an assumption that all distribution functions depend only upon the ratio N/Ω in this limit. Once the asymptotic form of the various diagrams is established, one can retain only the dominant diagrams to each order in λ. Doing this sometimes yields closed equations for the coefficients with few non-zero wave vectors. However, in order to get convergent results when terms of higher order in λ are included, one must again assume weak coupling, even when only dominant diagrams are included.

A more direct route to the final results of Prigogine is via the resolvent formalism, as developed by Resibois. To this end, define a resolvent operator R(z), corresponding to the Liouville operator, by

$$R(z) = (L - zI)^{-1} \qquad (12.44)$$

where z is a complex variable. From §9.3,

$$\rho(t) = e^{-iLt}\rho(o) \quad . \qquad (12.45)$$

Using the Cauchy formula for operators, eqn. (12.45) can be written as

$$\rho(t) = -\frac{1}{2\pi i}\oint_C \frac{e^{-izt}}{(L-zI)} dz\,\rho(o) \qquad (12.46)$$

where the contour C is shown in Fig. 12.3.

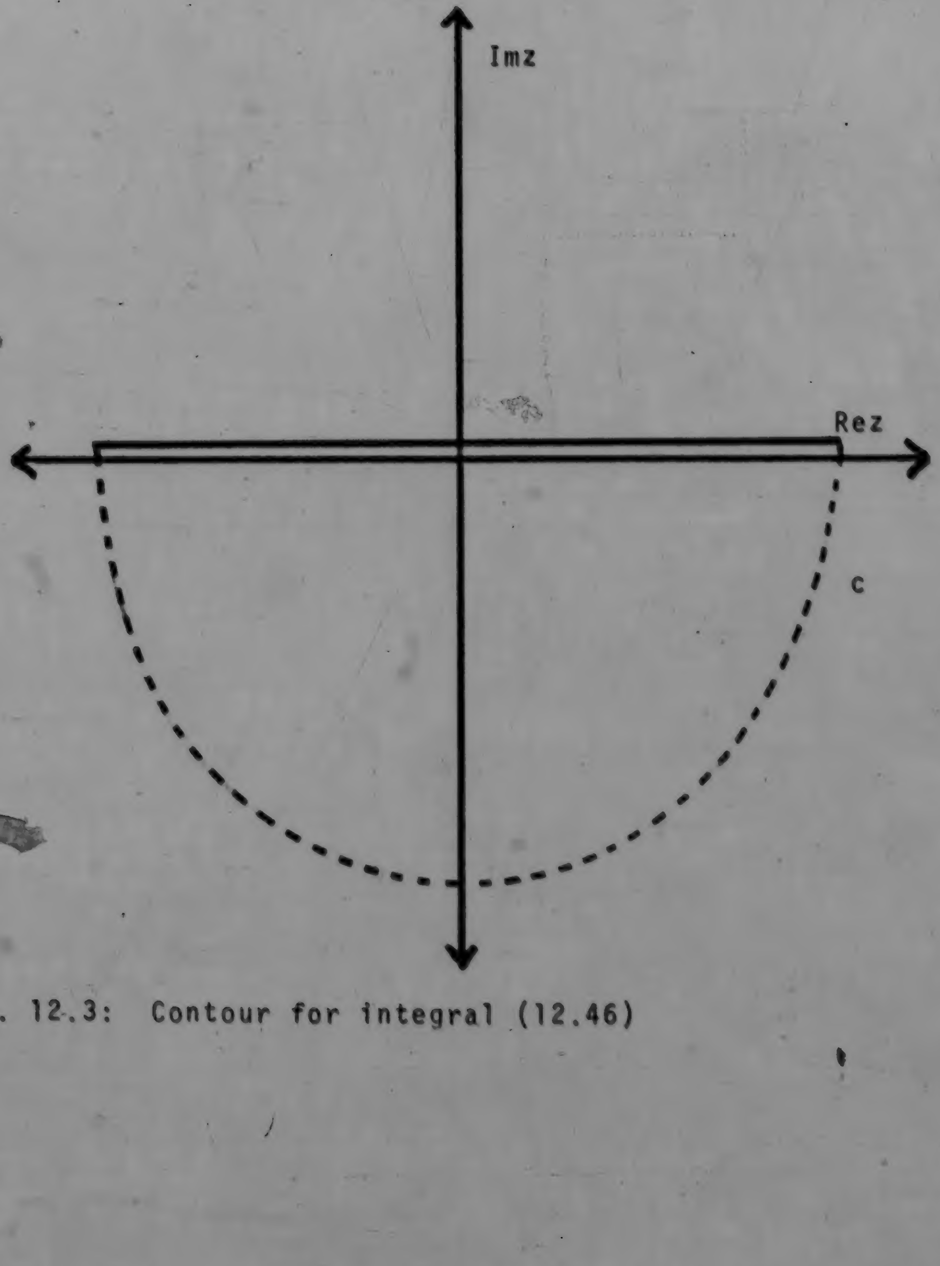


Fig. 12.3: Contour for integral (12.46)

Since $\mathcal{L}$ is Hermitian, all singularities of the integrand lie on the real axis and therefore the contour in Fig. 12.3 encloses all the singularities.

Using the operator identity

$$A^{-1}-B^{-1} = A^{-1}(B-A)B^{-1} \qquad (12.47)$$

for the perturbed and unperturbed resolvents:

$$A^{-1} = R(z) = (\mathcal{L}-zI)^{-1}$$

and

$$B^{-1} = R_o(z) = (\mathcal{L}_o-zI)^{-1} \quad , \qquad (12.48)$$

gives

$$R(z) = R_o(z) - \lambda R(z)\delta\mathcal{L}R_o(z) \quad . \qquad (12.49)$$

This can be iterated to yield the perturbation expansion

$$R(z) = \sum_{n=o}^{\infty} R_o(z)[-\lambda\delta\mathcal{L}R_o(z)]^n \quad . \qquad (12.50)$$

Next, Fourier expand $\rho(t)$ as

$$\rho(t) = \Omega^{-N} \sum_{\{\vec{k}\}} \rho_{\{\vec{k}\}}(t)\, e^{i\sum_{j=1}^{N} \vec{k}_j\cdot\vec{r}_j} \qquad (12.51)$$

where the $\rho_{\{\vec{k}\}}(t)$ in eqn. (12.51) are not to be confused with those in eqn. (12.35). Inserting eqns. (12.50) and (12.51) in eqn. (12.46) gives

$$\rho_{\{\vec{k}\}}(t) = -\frac{1}{2\pi i}\oint e^{-izt} \sum_{\{\vec{k}'\}} \sum_{n=o}^{\infty} <\{\vec{k}\}|R_o(z)[-\lambda\delta\mathcal{L}R_o(z)]^n|\{\vec{k}'\}> \rho_{\{\vec{k}'\}}(o)dz \qquad (12.52)$$

where the matrix elements of $\delta\mathcal{L}$ are given by eqn. (12.40) and those of $R_0(z)$ by

$$\langle\{\vec{k}\}|R_0(z)|\{\vec{k}'\}\rangle = \frac{1}{\sum_{j=1}^{N}\frac{\vec{k}_j\cdot\vec{p}_j}{m}-z}\prod_{j=1}^{N}\delta_{\vec{k}_j,\vec{k}'_j} \quad . \quad (12.53)$$

The matrix elements of $R_0(z)$ are referred to as propagating the correlation state between interactions.

For example, in the scattering situation described in §12.1, the second order effect of the scattering potential represented in eqn. (12.33) corresponds to the term

$$\langle o|R_0\delta\mathcal{L}R_0\delta\mathcal{L}R_0|o\rangle \qquad (12.54)$$

where the initial and final states have no correlations. This term is associated with the diagram of Fig. 12.4.

The expression (12.54) is said to represent the lowest order "vacuum-vacuum" transition.

For the vacuum state $|o\rangle$, $R_0(z)$ has a particularly simple form and the expression (12.54) can be written as

$$\begin{aligned}\langle o|R_0\delta\mathcal{L}R_0\delta\mathcal{L}R_0|o\rangle &= \frac{1}{z^2}\langle o|\delta\mathcal{L}R_0\delta\mathcal{L}|o\rangle \\ &= \frac{1}{z^2}\psi(z) \qquad (12.55)\end{aligned}$$

where the operator $\psi(z)$ defined by eqn. (12.55) is referred to as the finite frequency collision operator.

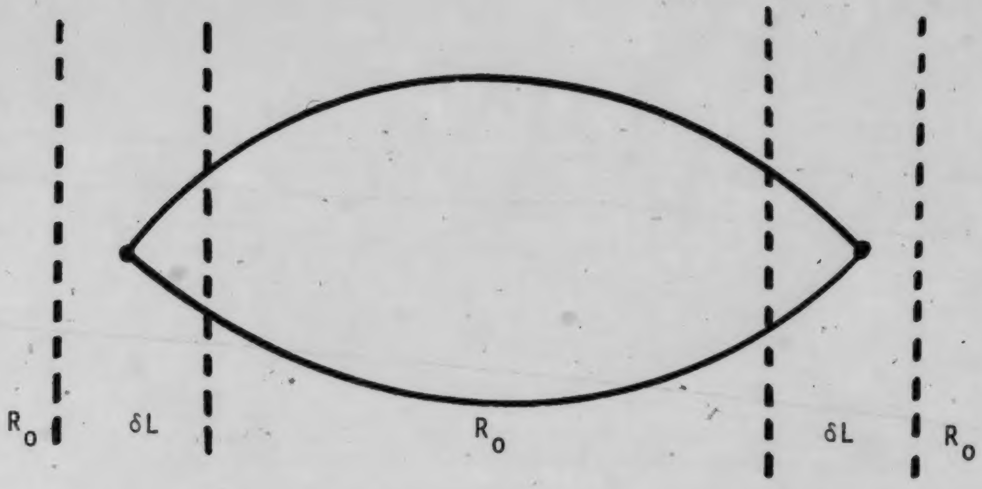

R_o
δL
R_o
δL
R_o

Fig. 12.4: Lowest order vacuum-vacuum diagram

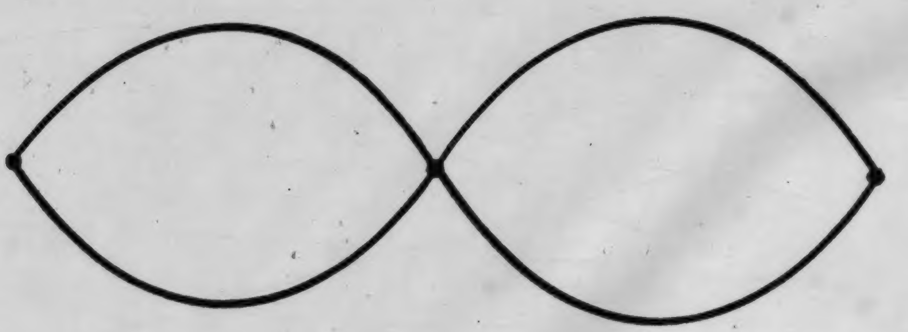


Fig. 12.5: Multiple collision diagram

The contribution of the term (12.54) to the evolution of $\rho_{\{\vec{k}\}}$ is given by

$$I = -\frac{1}{2\pi i} \oint_C e^{-izt} \frac{1}{z^2} \psi(z)\rho_0(\{\vec{p}\},t)dz \quad . \qquad (12.56)$$

Assuming that $\psi(z)$ is regular at z=0 and neglecting singularities of $\psi(z)$,

$$I = -\lambda^2 \frac{d}{dz}[e^{-izt}\psi(z)\rho_0(\{\vec{p}\},o)]\Big|_{z=0}$$

$$= \lambda^2[it\psi(o) + \frac{d\psi}{dz}\Big|_{z=0}]\rho_0(\{\vec{p}\},o) \quad . \qquad (12.57)$$

The first term of (12.57) grows linearly in time and is interpreted as a collision process. Thus, each of these vacuum-vacuum transitions, by introducing a double pole at z=0, describes a collision process. Similarly, a destruction or creation term gives a single pole at z=0 and so yields a contribution constant in time. A term of the form illustrated in Fig. 12.5 gives a triple pole and corresponds to two successive collisions. The terms ignored in eqn. (12.57), arising from singularities of $\psi(z)$ at $z\neq0$, will, in general, yield damped oscillatory terms.

To get a clearer understanding of these conclusions, it is worthwhile to study the structure of $\psi(z)$ in slightly more detail. From eqn. (12.55),

$$\psi(z) = \langle o|\delta \mathcal{L} R_o \delta \mathcal{L}|o\rangle$$

$$= \sum_{\vec{k}} \sum_{1\leq j<i\leq N} \langle o|\delta \mathcal{L}_{ij}|\vec{k},-\vec{k}\rangle\langle\vec{k},-\vec{k}|R_o|\vec{k},-\vec{k}\rangle\langle\vec{k},-\vec{k}|\delta \mathcal{L}_{ij}|o\rangle \qquad (12.58)$$

and

$$\langle\vec{k},-\vec{k}|R_o|\vec{k},-\vec{k}\rangle = \frac{1}{\vec{k}\cdot(\vec{v}_i-\vec{v}_j)-z} \quad , \qquad (12.59)$$

$$\langle o|\delta \mathcal{L}_{ij}|\vec{k},-\vec{k}\rangle = \Omega^{-1} V_{\vec{k}} \vec{k}\cdot\vec{D}_{ij} \quad . \qquad (12.60)$$

Converting the sum over $\vec{k}$ to an integral by using eqn. (12.34) gives

$$\psi(z) \sim \sum_{1\leq j<i\leq N} \vec{D}_{ij}\cdot\{\int \frac{|V_{\vec{k}}|^2 \vec{k}\vec{k} d^3\vec{k}}{\vec{k}\cdot(\vec{v}_i-\vec{v}_j)-z}\}\cdot\vec{D}_{ij} \quad . \qquad (12.61)$$

This implies a transition to a thermodynamically large system.

For an interparticle potential of range $1/a$, given by

$$V(\vec{r}) \sim e^{-ar} \qquad (12.62)$$

the Fourier components are

$$V_{\vec{k}} \sim \frac{1}{ik} [\frac{1}{(a-ik)^2} - \frac{1}{(a+ik)^2}] \quad . \qquad (12.63)$$

Thus, with this choice of potential, new poles appear in the integrand of eqn. (12.56). The contribution of these poles to I can be shown to be of the form

$$\frac{1}{(\vec{v}_i-\vec{v}_j)^2}\, e^{-a|\vec{v}_i-\vec{v}_j|t}\, \rho_0(\{\vec{p}\},o) \quad . \qquad (12.64)$$

In this expression, $a|\vec{v}_i-\vec{v}_j|$ is of the order of the inverse of the collision time for two particles, so the expression (12.64) is damped in a time of order of the collision time and is unimportant at large times.

Although it is not proven, it is assumed that the general terms in the expression (12.52) have a behavior similar to that of the term (12.55). If this is so, it is possible to draw rather general conclusions about the important terms in the general case. However, the assumptions which must be made restrict the validity of the conclusions. For example, in much of the work of Prigogine, Resibois, etc., it is assumed that there are no long range correlations or that the system interacts via essentially instantaneous collisions. Unfortunately, it is not clear what the explicit restrictions are upon the system in order to justify the very general assumptions which Prigogine makes.

In order to study the general evolution, the various terms are separated into creation, destruction, and diagonal terms. This separation is indicated by the study of simple terms such as (12.53) and is a formal one in that explicit demonstrations that such a separation is a fundamental one are lacking.

The general equation for ρ_o is

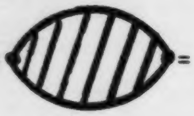
$$\rho_o(\{\vec{p}\},t) = -\frac{1}{2\pi i}\oint_C e^{-izt}\sum_{n=o}^{\infty}\{<o|R_o[-\lambda\delta\mathcal{L}R_o]^n|o>\rho_o(\{\vec{p}\},o)$$

$$+\sum_{\{\vec{k}'\}\neq o}<o|R_o[-\lambda\delta\mathcal{L}R_o]^n|\{\vec{k}'\}>\rho_{\{\vec{k}'\}}(\{\vec{p}\},o)\} \quad . \qquad (12.65)$$

This has been separated into a "diagonal" vacuum-vacuum term and a "destruction" term. Next, introduce the following operators and symbols:

(1) an irreducible diagonal diagram

$$-\frac{\psi(z)}{z} =$$

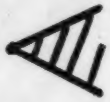


$$= -\frac{1}{z}\sum_{m=1}^{\infty}<o|-\lambda\delta\mathcal{L}[-R_o\lambda\delta\mathcal{L}]^m|o>_{irr} \qquad (12.66)$$

where "irr" indicates that there are no intermediate vacuum states, and

(2) an irreducible destruction diagram

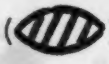
$$D_{\{\vec{k}'\}}(z) =$$



$$= \sum_{m=1}^{\infty}<o|[-\lambda\delta\mathcal{L}R_o]^m|\{\vec{k}'\}>_{irr} \quad . \qquad (12.67)$$

The complete diagonal contribution can be expressed in terms of the irreducible diagonal diagram as

$$\frac{1}{2\pi i}\oint_C \frac{e^{-izt}}{z}\sum_{n=o}^{\infty}($$

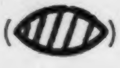


$$)^n \rho_o(\{\vec{p}\},o)dz \quad . \qquad (12.68)$$

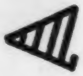
This is simply a reordering of the terms in eqn. (12.65) in a manner familiar from field theoretical perturbation theory.

Considering the complete destruction term allows one to write eqn. (12.65) as

$$\rho_o(\{\vec{p}\},t) = \frac{1}{2\pi i}\oint_C \frac{e^{-izt}}{z}\sum_{n=o}^{\infty} \Big($$



$$\Big)^n\{\rho_o(\{\vec{p}\},o) + \sum_{\{\vec{k}'\}\neq o}$$



$$\rho_{\{\vec{k}'\}}(\{\vec{p}\},o)\}dz \qquad (12.69)$$

$$= \frac{1}{2\pi i}\oint_C \frac{e^{-izt}}{z}\sum_{n=o}^{\infty}\left(\frac{\psi^+(z)}{-z}\right)^n\{\rho_o(\{\vec{p}\},o) + \hat{F}^+(z)\}dz \qquad (12.70)$$

where the abbreviation

$$\hat{F}^+(z) = \sum_{\{\vec{k}'\}\neq o} D_{\{\vec{k}'\}}(z)\rho_{\{\vec{k}'\}}(\{\vec{p}\},o) \qquad (12.71)$$

was introduced.

Differentiating eqn. (12.70) with respect to time gives

$$\frac{\partial\rho_o}{\partial t}(\{\vec{p}\},t) = -\frac{1}{2\pi}\oint_C e^{-izt}\sum_{n=o}^{\infty}\left(\frac{\psi^+(z)}{-z}\right)^n\{\rho_o(\{\vec{p}\},o) + \hat{F}^+(z)\}dz\ . \qquad (12.72)$$

If the Laplace transforms $F^+(t)$ and $\hat{\rho}_o(\{\vec{p}\},z)$ are defined by

$$F^+(t) = -\frac{1}{2\pi}\oint_C e^{-izt}\hat{F}^+(z)dz \qquad (12.73)$$

and $$\rho_o(\{\vec{p}\},t) = -\frac{1}{2\pi}\oint_C e^{-izt}\hat{\rho}(\{\vec{p}\},z)dz \qquad (12.74)$$

then

$$\frac{\partial\rho_0}{\partial t} = F^+(t) - \frac{i}{2\pi}\oint_C e^{-izt}\psi^+(z)\hat{\rho}_0(\{\vec{p}\},z)dz \ . \qquad (12.75)$$

Using the convolution theorem for Laplace transforms and denoting the Laplace transform of $\psi^+(z)$ by $G(t)$, this can be written as

$$\frac{\partial\rho_0(t)}{\partial t} = F^+(t) + \int_0^t G(\tau)\rho_0(t-\tau)d\tau \quad . \qquad (12.76)$$

This is the general evolution equation of the system. It is obviously non-Markoffian due to the presence of an integral term in eqn. (12.76). The term $F^+(t)$ can be said to represent the effect of the initial correlations on ρ_0. In simple cases, where there are no long range forces or correlations, it is anticipated that this term will decay in a time of the order of the collision time and thus can be neglected at large times.

The non-Markoffian integral term depends only upon the values of ρ_0 and represents the effect of the diagonal transitions with an arbitrary number of intermediate correlation states. Again, in simple cases, it is expected that $G(\tau)$ should be significant only for $\tau < t_{collision}$. In these "generalized Boltzmann situations", for $t >> t_{collision}$, the integral in eqn. (12.76) does not depend on all values of τ but only on those of the order of $t_{collision}$. Expanding $\rho_0(\{\vec{p}\},t-\tau)$ about t allows one to write an approximate Markoffian equation in place of eqn. (12.76) (cf. Prigogine [54]).

The evolution equation (12.76) is completely general and, as a result, is not a practical tool. In order to reduce it to a tractable equation, certain assumptions must be made. The first important assumption is that $F^{+}(t)$ can be neglected for large times. This is analagous to Bogolyubov's assumption that any initial correlations will be destroyed in a time of the order of the collision time and it restricts one to systems without "memory" effects.

The second important assumption often made is that the general expression in (12.76) can be reduced to a Markoffian form. This requires that the change in ρ_o due to transitions be negligible during a collision. This may be expected to hold only in systems in which a short, well-defined collision time exists, thus excluding systems with long range interactions.

It seems that, although the Prigogine theory formulates dynamics of large systems directly in terms of correlations and thus makes possible explicit formal statements of the various assumptions about correlations, the assumptions necessary to derive useful results are very similar to those of other kinetic theories. This somewhat detracts from the generality of the theory since it is generally very difficult to decide whether the assumptions made are satisfied in a given physical system.

As a final example of the Prigogine method, the case of a one-dimensional anharmonic crystal will be considered (cf. Prigogine [51]). This example is easily generalized to three dimensions. The system consists of N identical particles of mass m, spaced a distance "a" apart in the ground state. The potential energy associated with an arbitrary displacement of the particles from their ground state configuration may be written as

$$E=E_0+\frac{1}{2}\sum_{n,n'=1}^{N} A_{nn'}u_n u_{n'} + \frac{\lambda}{6}\sum_{n,n',n''=1}^{N} B_{nn'n''}u_n u_{n'} u_{n''} + \dots \tag{12.77}$$

where E_0 is the ground state energy and where the position x_n of the nth particle is

$$x_n = na + u_n \quad . \tag{12.78}$$

The cubic and higher order terms in eqn. (12.77) are important in describing the evolution of the system to equilibrium since the quadratic term yields only non-interacting normal modes (phonons).

The Hamiltonian of the system is

$$H = \sum_{n=1}^{N} \frac{(\dot{u}_n)^2}{2m} + (E-E_0) \quad . \tag{12.79}$$

In this case, instead of taking the free particle Hamiltonian as a starting point and regarding $E-E_0$ as the perturbation, it is useful to include the quadratic (harmonic) part of $(E-E_0)$ in H_0. That is, let

$$H_0 = \sum_{n=1}^{N} \frac{(\dot{u}_n)^2}{2m} + \frac{1}{2} \sum_{n,n'=1}^{N} A_{nn'} u_n u_{n'} \quad . \tag{12.80}$$

To find the corresponding L_0, use the well-known normal coordinate transformation

$$u_n = \sum_k q_k \, e^{ikna} \tag{12.81}$$

with the periodic boundary condition $u_{n+N}=u_n$. This restricts k to the values

$$k = m(\frac{2\pi}{Na}) \quad , \quad (m = -\frac{N}{2},\ldots,\frac{N}{2}) \quad . \tag{12.82}$$

If action-angle variables (J_k,α_k) are introduced, the Hamiltonian takes a simple form. The complex coordinates q_k are related to the action-angle variables by

$$q_k=(2Nm)^{-1/2}\{(J_k/\omega_k)^{1/2}e^{i\alpha_k}+(J_{-k}/\omega_{-k})^{1/2}e^{-i\alpha_{-k}})\} \tag{12.83}$$

and

$$\dot{q}_k = i\omega_k \, q_k \tag{12.84}$$

where $\omega_k(=\omega_{-k})$ is the frequency of the normal mode k, defined by

$$m\omega_k^2 = \sum_{n,n'=1}^{N} A_{nn'} \, e^{ik(n-n')a} \quad . \tag{12.85}$$

With these definitions,

$$H_o = \sum_k \omega_k J_k \tag{12.86}$$

which is the expected form for a collection of independent harmonic oscillators.

The perturbation which will be considered is the cubic term in eqn. (12.77). In terms of action-angle variables, this can be shown to be

$$\lambda V = \sum_{k,k',k''} \sum_{\varepsilon,\varepsilon',\varepsilon''=\pm 1} V_{\varepsilon k,\varepsilon' k',\varepsilon'' k''} \left(\frac{J_k J_{k'} J_{k''}}{\omega_k \omega_{k'} \omega_{k''}}\right)^{1/2} e^{i(\varepsilon\alpha_k + \varepsilon'\alpha_{k'} + \varepsilon''\alpha_{k''})} \tag{12.87}$$

where the term $V_{k,k',k''}$ is given by

$$V_{k,k',k''} = (2Nm)^{-3/2} \sum_{n,n',n''=1}^{N} \frac{1}{6} B_{nn'n''} \, e^{i(kn+k'n+k''n'')a} \, . \tag{12.88}$$

It is straightforward to recast the Liouville operator in terms of normal modes and action-angle variables (cf. Prigogine [51]). Doing so gives

$$\mathcal{L}_o = -i \sum_k \omega_k \frac{\partial}{\partial \alpha_k} \tag{12.89}$$

and

$$\mathcal{L} = i \sum_{k,k',k''} \sum_{\varepsilon,\varepsilon',\varepsilon''=\pm 1} V_{\varepsilon k,\varepsilon' k',\varepsilon'' k''} \left(\frac{J_k J_{k'} J_{k''}}{\omega_k \omega_{k'} \omega_{k''}}\right)^{1/2} e^{i(\varepsilon\alpha_k + \varepsilon'\alpha_{k'} + \varepsilon''\alpha_{k''})} \left\{\frac{1}{2}\left(\frac{1}{J_k}\frac{\partial}{\partial \alpha_k} + \frac{1}{J_{k'}}\frac{\partial}{\partial \alpha_{k'}} + \frac{1}{J_{k''}}\frac{\partial}{\partial \alpha_{k''}}\right) - i\left(\varepsilon\frac{\partial}{\partial J_k} + \varepsilon'\frac{\partial}{\partial J_{k'}} + \varepsilon''\frac{\partial}{\partial J_{k''}}\right)\right\}. \tag{12.90}$$

The matrix elements of δL are

$$\langle\{n\}|\delta L|\{n'\}\rangle=\frac{1}{(2\pi)^N}\int_0^{2\pi}\cdots\int_0^{2\pi} e^{-i\sum_j n_j\alpha_j}\,\delta L\, e^{i\sum_j n'_j\alpha_j}\, d\alpha_1\cdots d\alpha_N \quad . \tag{12.91}$$

The sets {n} denote the occupation numbers of the various normal modes. The only non-zero matrix elements are of the form

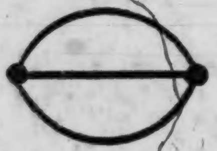
$$\langle\{n\};n_k,n_{k'},n_{k''}|\delta L|\{n\};n_k+\varepsilon,n_{k''}+\varepsilon',n_{k''}+\varepsilon''\rangle \quad . \tag{12.92}$$

The matrix element (12.92) is referred to as a three phonon interaction.

The equations for the $\rho_{\{k\}}$ can be solved by iteration, as in earlier discussions. If, in the weak coupling limit, only the lowest order non-vanishing terms are kept in the equation for ρ_o, then the important contribution is

$$i\psi_2^+(io) \equiv \langle o|\delta L R_o \delta L|o\rangle \quad . \tag{12.93}$$

This may be represented by the "three phonon vacuum-vacuum diagram"



In this approximation, the equation for ρ_o is

$$\frac{\partial\rho_o(t)}{\partial t} = i\lambda^2\psi_2^+(io)\rho_o(t) \quad . \tag{12.94}$$

This is a Boltzmann equation for phonons and was first derived by Peierls (1929) using arguments similar to those Boltzmann used for gases. It follows from the general evolution equation (12.76) by dropping all destruction terms and non-Markoffian effects and can only be valid for weak anharmonicities or short times. It describes an irreversible evolution to equilibrium and an H-theorem can be derived from it.

This concludes the discussion of the Prigogine theory of irreversible processes. Further examples and discussions can be found in the references (Prigogine [51] and [54], see also Jancel [42]).

BIBLIOGRAPHY

[1] Glansdorff, P. and Prigogine, I., Thermodynamic Theory of Structure, Stability and Fluctuations (Wiley-Interscience, London, 1971)

[2] Hofelich, F., Z. Physik 226 (1969) 395

[3] Meixner, J., Critical Review of Thermodynamics, Ed. Stuart, Gal-Or, and Brainard (Mono Book Corp., Baltimore, 1970)

[4] de Groot, S. R., and Mazur, P., Non-equilibrium Thermodynamics (North-Holland, Amsterdam, 1962)

[5] Frenkel, I., Kinetic Theory of Liquids (Clarendon, Oxford, 1946)

[6] Lodge, A., Elastic Liquids (Academic Press, London, 1964)

[7] Miller, D. G., Chem. Rev. 60 (1960) 15

[8] Hoffman, K., and Kunze, R., Linear Algebra, Second Edition (Prentice-Hall, Englewood Cliffs, New Jersey, 1971)

[9] Truesdell, C., Rational Thermodynamics (McGraw-Hill, New York, 1969)

[10] Hooyman, G. J., and de Groot, S. R., Physica 21 (1955) 73

[11] Davies, R. O., Physica 18 (1952) 182

[12] Hooyman, G. J., de Groot, S. R., and Mazur, P., Physica 21 (1955) 360

[13] Casimir, H. B. G., Rev. Mod. Phys. 17, no. 2 (1945) 343

[14] Onsager, L., Phys. Rev. 37 (1931) 405

[15] Glansdorff,P., and Prigogine, I., Physica 20 (1954) 773

[16] Schlögl, F., Z. Physik 243 (1971) 303

[17] Schlögl, F., Z. Physik 248 (1971) 446

[18] Bessenrodt, R., Z. Physik 264 (1973) 171

[19] Landau, L. D., and Lifshitz, E. M., Statistical Physics (Addison-Wesley, Reading, Mass., 1958)

[20] Schlögl, F., Z. Physik 253 (1972) 147

[21] Chandrasekhar, S., Hydrodynamic and Hydromagnetic Stability (Clarendon, Oxford, 1961)

[22] Whitehead, J. A., Fluctuations, Instabilities, and Phase Transitions (Plenum, New York, 1975) pp. 153-161

[23] Panchev, S., Random Functions and Turbulence (Pergamon, Oxford, 1971)

[24] Uberoi, M. S., Phys. Fluids 6 (1963) 8

[25] Isihara, A., Statistical Physics (Academic Press, New York, 1971)

[26] ter Haar, D., Rep. Prog. Phys. 24 (1961) 304

[27] Bogolyubov, N. N., in Studies in Statistical Mechanics, de Boer and Uhlenbeck, eds. (North Holland, Amsterdam, 1962)

[28] Trofimenkoff, P. N., and Kreuzer, H. J., Am. J. Phys. 41 (1973) 292

[29] Ziman, J. M., Principles of the Theory of Solids (Cambridge University Press, Cambridge, 1965)

[30] Cohen, E. G. D., Physica 27 (1961) 163 and
Physics 28 (1962) 1025

[31] Green, H. S., The Molecular Theory of Fluids (Dover, New York, 1969)

[32] Terletskii, Ya. P., Statistical Physics (North Holland, Amsterdam, 1971)

[33] Kestin, J., and Dorfman, J. R., A Course in Statistical Thermodynamics (Academic Press, New York, 1971)

[34] Wu, Ta-You, Kinetic Equations of Gases and Plasmas (Addison Wesley, Reading, Massachusetts, 1966)

[35] Blatt, J. M., Prog. Theor. Phys. 22 (1959) 745

[36] Hobson, A., Am. J. Phys. 34 (1966) 411 and Phys. Lett. 26A (1968) 649

[37] Lee, C. T., Can. J. Phys. 52 (1974) 1139

[38] Kreuzer, H. J., and Teshima, R., Can. J. Phys. (in press)

[39] Katz, A., Nuovo Cimento 33 (1964) 97

[40] Huang, K., Statistical Mechanics (John Wiley & Sons, New York, 1963)

[41] Hahn, E. L., Phys. Today 6 11 (1953) 4

[42] Jancel, R., Foundations of Classical and Quantum Statistical Mechanics (Pergamon, London, 1969)

[43] Kubo, R., J. Phys. Soc. Jpn. 12 6(1957) 570

[44] Nakano, H., Prog. Theor. Phys. 15 (1956) 77

[45] Bernard, W., and Callen, H. B., Rev. Mod. Phys. 31 4(1959) 1017

[46] van Kampen, N. G., Physica Norvegica 5 3-4(1971) 1279

[47] Nyquist, H., Phys. Rev. 32 (1928) 110

[48] Callen, H. B., and Welton, T. A., Phys. Rev. 83 (1951) 34

[49] Prigogine, I., and Balescu, R., Physica 25 (1959) 281

[50] Prigogine, I., and Balescu, R., Physica 25 (1959) 302

[51] Prigogine, I., Non-equilibrium Statistical Mechanics (Interscience, New York, 1962)

[52] Resibois, P., Physica 27 (1961) 541

[53] Wu, Ta-You, Theory of Irreversible Processes in Neutral and Ionized Gases, Lecture Notes, University Of Lausanne (1960-1961)

[54] Prigogine, I., in Topics in Non-Linear Physics (Proceedings of the Physics Section of the International School of Non-Linear Mathematics and Physics - Munich 1966, ed. N. J. Zambusky, Springer-Verlag, New York, 1968)

[55] Kreuzer, H. J., and Kurihara, Y. (in press)